

> O <
OI IO Intelligence

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 09972467-2-509-578-ab037733.res made by jdelaval on Thu 20 Jun 102 6:17:05

Query sequence being compared: US-09-972-467-2 (509-578)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of US-09-972-467-2 (509-578) with:
File: ab037733.seq

100-
-
N -
U - 50-
-
-
-
O -
F 10-
-
S -
E 5-
Q -
U -
N -
C -
E -
S 0-
-
SCORE 0 1 8 16 23 31 39 47 54 62 70
STDEV 0 1 1

Similarity matrix Unitary 6 K-tuple 2
Translation Frame 1
Mismatch penalty 1.00 Joining penalty 20
Gap penalty 0.05 Window size 32
Off score 0
domination group 0

SEARCH STATISTICS

Scores: Mean 16 Median 7 Standard Deviation 26.06
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 10271
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Init. Opt.

Sequence Name Description Length Score Score Sig. Frame
1. ab037733 TOIG of: ab037733 check: 480 1712 70 70 2.07 2

The list of other best scores is:

Sequence Name Description Length Score Score Sig. Frame
2. ab037733 TOIG of: ab037733 check: 480 1711 7 14 -0.35 6
3. ab037733 TOIG of: ab037733 check: 480 1712 7 12 -0.35 4
4. ab037733 TOIG of: ab037733 check: 480 1711 6 13 -0.38 5
5. ab037733 TOIG of: ab037733 check: 480 1713 6 12 -0.38 1
6. ab037733 TOIG of: ab037733 check: 480 1712 5 11 -0.42 3

1. US-09-972-467-2 (509-578)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:pbuescriptII SK plus
clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M., Ohara,O.,
TITLE Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res 7 (1), 65-73 (2000)

MEDLINE
JOURNAL 20181126
REFERENCE 2 (bases 1 to 5139)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
297-0812, Japan (E-mail:cdna@nfo.kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)

FEATURES
source
Location/Qualifiers
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/issue_type="brain"
/clone_lib="pbuescriptII SK plus"
1..4417
/gene="KIAA1312"
1..4417
/gene="KIAA1312"
/note="Start codon is not identified."
/product="KIAA1312 protein"
/protein_id="BA92550.1"
/db_xref="GI:7242979"

translation="NSEHTAVISLCSGMIGTFRSHDGYFIPLQSMDEDEDEBONK
PHILYRRSAPQREPTGRIACDTSERKNSHSDKKTARKKGERINAGDAALNSG
LATEASAYGNKTDNTRERKTRKRLFLSPYEVIVADNRWVSYHGEVLAJYIL
TMSIASIYKDSIGNLINTIYVNLVIHNDODSPSIFENNQTILKNECOMSHKNS
PGIHHDTAVALLRQDICRAHDKDFTLGLAELGTICDPYRSCSISDSGISTAPAKS
ELGHVNMNPMDDNNKCKEESVSRQHYVAPITNFTTNPMMSKSKRTITTELDITGYG
ECLINPESERPYPLPYOLPGIILYNWKOCELIFGPSSOVCPTMOCRRIMCWNVGVH
KGGRTDHTWADGTCEBGRHCKYGVCEVPEKMDVPTVDSGMSNPSFGSCRTGCGI
KTAIRCNKPEPKNGKCYCVRMRKFKSCNTEPCIKRQKDFDEQCAFHGHENING
LIPNVRWPKYSGIIMKDRCKLFCRVAGNTAYAYOLRDYIDGTGCGOTINDICVGGIC
RQAGCDHVINSKARDKCGVCGDNNSSCKTVAGTFNTYVHYGNTVYRIIPACATINDIVR

KTAIRBENPREPNKNGKXYCYGARMKFKSCNTPOLKROPRDRESCDCAHPGFKHKNH
 LLAIPREPNKPYSCGLIMEDROCKLEPCRYAGNANAYUOLPRVIDGTSCODTIDICVQIGL
 ROAGCDHYLNSKARDDKCYCGGDNSSCKTAYGTFNTYHGTITNTAYIPMATIIVRA
 QHSGSEGETDDNTLALSSKGEFLNENFTYKMAKREIGNAWVEYSGSTAEVRIEN
 STHIEBELLOLVYVGRKLVNPDVRYFNLETERPOOFYTNAMWVOSACSPKQGBER
 KRKYLCYRESBDDQVYDRCEDRLPQPHITPESCTGDDLRNHNARSRECSACSLGTR
 TLDYCAKYSRDLTSTOKEDVDFGSCSHPRKSNRCKSGEECTYGMGYSAMEVCEKSD
 GGTORRAICIVNTRNDVLDJSCKTHOEKVIORCSPPCKMGMSCELYTQCKSH
 KHRVWOCFGEDRLNDMPCEBETKPTSMOICOQCEPSCWMOGSPMGSCSVTCOGYOLH
 AVKCIIGTYSVVDNDCNATRTPTDQDLPCHPHRAPEARESTRYSAPBOWMEF
 GSWTPCSATCGKTRMRYVSCRDNGSVDBASACATLPRAPEAKESVYTPGKORML
 WSSCSVTGCGGGRATROYAVYSDHVIJDSRECDQDITPEIDQDSMBECPORPTDGY
 AOHPEONEDYRPRRSAPRTHVYLGNOVIRGCPMCACTSGAGSSORRYVUNOQPNQY
 ANDOVERIKEDQACSESGCPQMAVYENMBGCTILCGGJITRLVYVQORSNGEFPDIL
 SCLEILKPPDRBCNTIACPHDAWMASTGPMSCVSGCRGVGROKNOVIMADVSHLS
 DYCHLAKPHGRKCGRCCKMAWAGMASOCVSGCRGVGROKNOVIGTHK IARETHE
 CNPTRTRESREDCDQGPCRPLTYWRAEEMOEDCTKYSRYKRVVCDVDNKNNEVHGAR
 CDSKRYREDRSCSLQCELYYVWITGMESEVPSWEL"

```

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
      Cital Score = 7 Optimized Score = 14 Significance = -0.35
      Residue Identity = 21% Matches = 20 Mismatches = 46
      Gaps = 29 Conservative Substitutions = 0
      Translation Frame= 6

```

IGRWASSLPBSCHTPSTHPSSTLPMLSXWMLQFPGRYLXFPCSDPDILAGIQTAAFTVHRIALVHTACPMH	210	220	230	240	250	260	270	280
LVEVVIHHTHAAVGAFAAGLGAFLPRLCIPSPVEGTPMALAFAFRJSGSWWALGLSGFYVCSDLTAMRPLL	290	300	310	320	330	340	350	
YASATGHRALTPSAPFFPGASSSSALSVPMMLISQVLVTTEFKMASIECHAVNLSLEMPSTTHRTTRABRAST	360	370	380	390	400	410	420	
PCSVVWTSMCVTLTLTIGREIKNFQAIWKPRTGPTLTDHOSCVASTQTLASLPAISLPTETRAGFTGSL	430	440	450	460	470	480	490	
IRFYSLHTVWGVSILFILTLYNNPRLGASITGTCTSPGASSPLVTEHMGAGAGAPGTIVYLEGVLSXAT	500	510	520	530	540	550	560	
WVGPLRAXHGTVLVSFMDIILVTLTPRIDHVTVTVDTNHLPGCPPLPTGHRARAPVOGILPTFGCHRTFFELC	570	580	590	600	610	620	630	640
HMSRSGSTGTLVSHRALILISADVSHRGPPTSGXAMGPRPKSPGLGSMWCTVOCASPRFSGMCRMTXKXFTVL	650	660	670	680	690	700	710	
GISWSSCCITVIVYKHHSIPNDAFHCSXVLSLSTSDTALSPTRLPGCTFFRLLSLHRCMLGLRVHSITIX	720	730	740	750	760	770	780	
SIFTKLPDLAVLMPFSTGDOALXPVSRFLMTRRELTAPLNGNLFMLCMAFAVIOYIISGIDNSPSPGLPTVY	790	800	810	820	830	840	850	
TAEXFSPBRIAPTRATVTPKXAFPTVAMFGMAAKTILNLSLPIOPALFETVDVQCAVTOATLGTFPTPAG	860	870	880	890	900	910	920	
NMPPOVYVCTTGESVNSRLGOPLASLIRNSKLLIRFPGANKFSFPLLAGFTACLEPMTVVPKLLIRLTENN	930	940	950	960	970	980	990	1000

IERIVAYIGVQVLSHQRLKQKRELLNACVWNSFSGLSGPTVLVHSIPNANFPGHGDKNKVSIXOEPETATXAX
 1010 1020 1030 1040 1050 1060 1070
 SXVVYVICFPXETVLPINIGSTSWNSDHSIYTIMVCICKSCCHCFAXRIIATNTNFTLSPGFXVXNMIASSL
 1080 1090 1100 1110 1120 1130 1140
 PAKALDTDIICVLATRSSIYHSVSKLIVGCVPHCSAEQLAPVLHONSTVFRDPAHIGKOTVDKMLPKVAST
 1150 1160 1170 1180 1190 1200 1210
 X
 QLP-----
 I I

```

10      20      30      40      50      60
GIIYVNNKOCF-----LINGSOYCPTMOCRLKMCNNV---NGYKGCRTQHTTPADSTEEBEPKHC-
LFISEVSLLEFLMWLRVAGLKCHSTS---YTFSEITJFMFWSALNSGCEADSTCSGASKOTPTSPISICH
1220    1230    1240    1250    1260    1270    1280

-----KYGFVY      X
      |      |
GDJHFFGNTKSLILAYLSLRLAPRPGCVSLSPAAYFVYSIDVJAPRPSALHHIHWHLRTWSKNQOTFLFTHVYK
1290    1300    1310    1320    1330    1340    1350

DAWOLDROHVGSGFRFVYKQTLALINSYKXKLSDJESTLLXPHRGVGEVOCMSHDMJGLTINSEFFTEFVVUIMRH
1360    1370    1380    1390    1400    1410    1420

VHMMQMLMDKRSCTOSTITFFNNRRAISGITRNGSOFSOAKXITTFVSSADILSCXENSIAMDSTWTYVLRMP
1430    1440    1450    1460    1470    1480    1490    1500

LAKVFECCLSTIRKGRPILFLMHNHYVNNHNINXISNTWFEIDRGYNXKXSXNIVLKYFSAVXNHSVVCCHQO
1510    1520    1530    1540    1550    1560    1570

DFYKPMWIGKKTFECPSGVGLFSCVYRILITISRKCLCCXAAAXCCVYVNSOVNNPSPSCSGLFVETVPYLFVFX
1580    1590    1600    1610    1620    1630    1640

GYTCMPCSCXGSLSGAABINDVGEVLEFLFLEFLTHRLXWENKIIPITMRPCCAQHSXAERDGRVIG
1650    1660    1670    1680    1690    1700    1710

```

3. US-09-972-467-2 (509-578)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5135

LOCUS	AB037773	5139 bp	linear	PRI 14-MAR-2000
DEFINITION	Homo sapiens mRNA for KIAA1312 protein, partial cds.			
ACCESSION	AB037773			
VERSION	AB037773.1	GI:7242978		
KEYWORDS				
SOURCE	Homo sapiens brain cDNA to mRNA, clone_lib:pbLuscriptII SK plus			

REFERENCE	1 (sites)
AUTHORS	Nagase, T., Kikuno, R., Ishikawa, K. I., Hirosewa, M. and Ohara, O.
TITLE	Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro
JOURNAL	DNA Res. 7 (1), 65-73 (2000)
MEDLINE	20181126

AUTHORS Ohara, O., Nagase, T., and Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, 1520-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: oohara@kazusa.or.jp, URL: <http://www.kazusa.or.jp/hnuge/>, Tel: +81 436-52-3913, Fax: +81-436-52-3914)

FEATURES	
source	location/qualities
gene	1..5139 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="fh11767" /tissue_type="brain" /clone_lib="pb BluescriptII SK plus" 1..4417 /gene="K1AA1312"
CDS	<1..4417 /gene="K1AA1312" /note="Start codon is not identified." /codon_start=2 /product="K1AA1312 protein" /protein_id="BAA92550.1"

```

/db_xref="GI:7242979"
/translation="NSEHTAVISLCSGMLGTERSHDGVFEIEPLQSDQEDDEEONK
PHIIRRSAPORREPTSGRHACDTESEKHNKSKKKTRAKNGERINLAGDAALNSG
LATEAFSANGKTDNTREKTRHRTREFLSYPRFVULVADNRWYSYGENLOHYTIL
TLMSTIASITKDPISIGNLINIYVNIIVIHNEBDGSIENMAOTILKNPCQONKSN
PGGIHDTAVLITRQDICRAHDCDITGLAELCTIDIPRSCSISDSLSLAFITAH
ELGHVNMPPHDDNNKCEBVGSPQHVMATPLNFYINPMWMSCKSKRYITFELDYG
ECLLNBPESRPPLPVOLPGILYNVKOCCELLFPGSQVCYPMQCRILMNQVNH
KGRTOHTPADTGECEPGHCKYGFVCYKREMDVPVTDGSMGNSPFGTSCGGGI
KTAIRECNREPEKNGKCYGCRMRKFSCTNEPCIKOKRDEPCAHGDKHFNING
LLPNWVWPKYSIILKDKCKLFCRAGNTATYQLDRVIDTGPCODNDICVGLC
RQAGCHVLNLSKARNDKCVGCGDNSSCTVAGTFTVIGINTVIRTAGATNDIVR
OHSFSETDDNTLALSSSGEFLNGNVTMAKEIIGAVVEISSEIAVRIN
STDRIDELLQVLSVGLKLYNDVRYSPNIPIEDKQOQYKMHVSHGFWQASPCOGER
KRLVCTRESDDQLVSDORDLPQGHITPEPGTCDLRMHVASBEGSACCGGER
TLDIYCAISRLDGKTEKYGSCSHSPNENKSGCMTNGMYSWTEGSCSD
GTOBRRATCIVNRNDVLDSSKCTHOEKTORSPFCPOKSGDMSCLATYCKGKH
KHQVWCQOGEEDLNDKMDPEKPTSMOTCOQPEASNOAGPWCSTYCGOQOLR
AVACITIGTMSVVDNDCNATRPTDIOCELSHPRAAETRTSTYSAKRTQRF
GSMTPCSATGCKGTMRVYSCDENGSVADSEACALPRPVAKEDSCVTPCQOMALD
WSSCSVTCGGGRATQVMCVNSDHYIDRECDODYIPEITDSCSPCOPRTDPSGL
AOHPONEDYRPRASPSRTHVLGDMORTGPGACSTGACGSOBRVYVCCDENGYT
ANDCEERIKPEDEORACEGSPQWAGYGMNBECKILGGGIFTRILYVCOMNBERPDL
SCCELLKPPREOCNTHACBDAANSTGSSVSCSGRHHQRVNYCAKAGSHLES
DYCKHLAKPHGKCKGRCGPKKAGAMSCSYSCGRYQOHVGOJQITKRLTARETE
CNPTYTRERDCQPRCLPTMYRAEWEQCTKTCGEGSYRKRKVVYVDNKNKEVGAR
CDVSKRPVRESLOPCFCEVWITGEMSEVPSWEL"

```

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803

Initial Score = 7 Optimized Score = 12 Significance = -0.35
 Residue Identity = 20% Matches = 16 Mismatches = 40
 Gaps = 24 Conservative Substitutions = 0
 Translation Frame = 4

```

LCVPTXQPTCLCTPRPDTHXLOAFPHLGHLPNHFILCPGCLACLOXLSLKMPLSPAMOXTERCLCPR
350 360 370 380 390 400 410
POETEHEHGLFVLAHSAAGQACVLHCSNGSLRISOLKSGNRSPDLRXQTTSLVLPHPHSLVSPQLPXA
420 430 440 450 460 470 480
HXGGPDSQALCSGGFLLSTQSFAYVYPPSSXHTTRMDPQAQVLEHAPOGVLHMFPESTWRLGLALNGR
490 500 510 520 530 540 550
XSSFWKGCXAKPLSLGVLXGHGDMQMSVSGIXSMHSLSRITSXLTHITCRVALPCQVTEQELQSKAF
560 570 580 590 600 610 620
HCPQGVTEHSSFAIGLGRVQALSSATEPPSSKQTLTILVPPFQVAHEGVODPNRMWYVYALVLLYSGA
630 640 650 660 670 680 690 700
AGGGXHDGNSQSWVSLVLAALQSLSTTDIXVPMHFTALSMWPCPOVTLHCPQGPACODAHSGCXQVCID
710 720 730 740 750 760 770
VGLVSGHILLSFNSSPRXHQTWRCCLPFPQVYTKHSDQSPDFHCGGNSLHLKMVTFESXCVHLLSSSTPRV
780 790 800 810 820 830 840
LTQJALLMLWPPSQLFEHSHVQAEKRPVVLHSPHFSSRLGLGLMLQKPSSTFVSPPSLLYLAQXMSVNX
850 860 870 880 890 900 910
PKPIMALHSLLLATLACHLRSGSVPGQSVWCPGWSRSHLSSETYSXSDSLVOTSFRRFSPWQGLHACHGPXL
920 930 940 950 960 970 980
FOXNKGSLSSIGILKEKXRTSGLYNFPDXTKXSSCSMRSEVELLSTAVSDPLYSTFAFPMRISLAIYTTK
990 1000 1010 1020 1030 1040 1050 1060
FFPSKNSPLLLDKAKXLSSSVSPEKLCGRSILVAPAGITTYLXPKACTYLANPATVYHEELSPOTPHSL
1070 1080 1090 1100 1110 1120 1130
RAPEFKTXSHPACRORPWTQISFVSNPQGVPSITLSRSXXAVPATQLQNNLHRSFRIRPLVLYGTORTGSR
1140 1150 1160 1170 1180 1190 1200

```

```

PLMKCFPSKXAHCCSSSRKSRFCILRHGVSLODINFILRPTQYPPFEGSGLHRSMAVLMPPHYLEQVPKGL
1210 1220 1230 1240 1250 1260 1270
QLPQDSVYTGTSISLGTQNPVYLQCFPGSHSVPSAOGVCXVRPQLTPLTLHQSRHLGCIYGHTEPQPKIN
1280 1290 1300 1310 1320 1330 1340 X
VKNQCLILGPGSQVCYPMQCCRILMCNNVNGVHKGRTQHP-----WADTGECEGKCKYGFV
1350 1360 1370 1380 1390 1400
SHCLFTLXMPG-----WTGKXGLDGSLSKSPXPVSKNSYIFRLHFDHIOGLVXK--FSV
1410 1420 1430 1440 1450 1460 1470
GAMTCGLTTPSSLHLLSSXGMLNTWPSMAIVKAVLNPSSLIEQLXGSONVPSSARPKVSHLSXALO
1480 1490 1500 1510 1520 1530 1540 1550
SCLVKTAVSXWIPRGLFPECCHMOWFEFNVVXALDMEGPSCSLXITIKETITLILKEPILGSLXIEATDI
1560 1570 1580 1590 1600 1610 1620
KVKIXCRPSWYETILLSATTKTSTNRGXDKRRVILWVLFSLVLSVLYLXAVNASVAKPLFNATSPARL
1630 1640 1650 1660 1670 1680 1690
ILSPHLALVFLSLCLFLCSEVSHACLPVEGSLMGALRLXMMNGLFCSSSSSCSIDCSGSIKXSPXDR
1700 1710
NPSIPEORLMTAVCSE

```

4. US-09-972-467-2 (509-578)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 AB037733
 VERSION AB037733.1 GI:7242978

KEYWORDS Homo sapiens brain cDNA to mRNA, clone_11b:pb1uescriptII SK plus
 clone:fh11767.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (sites)
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro

DNA Res. 7 (1), 65-73 (2000)

JOURNAL

2 (bases 1 to 5139)
 Ohara,O., Nagase,T. and Kikuno,R.
 Direct Submission
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)

FEATURES

SOURCE

1. 5139
 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="fh11767"

/tissue-type="brain"

/clone_id="pb1uescriptII SK plus"

1. .4417

/gene="KIAA1312"

<1. .4417

/gene="KIAA1312"

/note="Start codon is not identified."


```

/codon_start=2
/product="KIAA1312 protein"
/protein_id="BA92550.1"
/db_xref="GI:7242979"
/translation="NSEHTAVISLCSGMLGTFRSHDGYETIEPLQSMDEDEEONK
PHIYRSAPQREPSTGRHACDTSSEHKRNRKDKKTRARRKMGRIINAGDVALLNG
LATEFASVGNKTDNTRKTRHRRKTRFSLPREVEVLVADNMYVHGHNLOHTL
TMSIVASITYKDPSTGNINIVINLVLIHNEODPSTSEFAOTTLLKPNOMOSKNS
PGGHHIDTAVLLTRKODICRAHDCDTGLABELGTICDPYRSCSISESGSLATTIH
ELGAFENRPHDNNKCKEEGVKSPQHWAPLNTTYNEMWMSKSRKYITIEFLDTG
ECLINPEPSRPLPYVLPGLILVNNKCELIFFGQVQCYMMCKRMLNNVNHG
KGCRTQHPADGTECEBPCKHCKYGFVCKEMADYVLDGWSGSPGTSRCSGGI
KTAIRECNRPEPNKGGKCYGVRMKFKSCNTEPKDQDRDRDSCAHFQKHNING
LLPNRWPKYSGILMDRCKLFCRVAGNTAYOLDRIDGTCGGDINDICVQGLC
ROAGCDHYLNKARDKCGCGGSDNSCKYVAGTGFNTYHYVNTVTRPGATIDVR
OHSFEGEDDDVTLLSSKGEFLINGVYTMAREIRINAYEVSSESTAYERN
STRIEQLQLQVLSVGLYNPDVATSNITIEDPQOFTNNSGPMWACKPQOGER
KRKLVCRESQDLVSDQCDRLPQPGHITPECCDILRNHVASRSECAQCLGTR
TLDIYCAKYSRLDKTEKVDGFCSSHPRKPNREKSGECNTGWMRYSAWTECSKD
GOTORRAICVNTRNDVLDSDSKCTHOEKVITIQRCSEFPCQMKSGDSECLVTCGH
KHROVCOFGEEDRLDMCDPEETKPTSMOTCOOPFCASMOAGPWGSCVTCGGYOLR
AYKCIITGYMSYVDNDCNATRPDTODCELPSGHPPAPAPETRRSTYSAPRTOMR
GSMTPCSATCGKCTRMRYVSCRDENGSAVDSACTTLPRAKEECSVTPCGOMKALD
WSSCSVTCGGQKATQVACVNTSDHVIDRSECDQDYTEIDQDSMSFCPORTPSGL
AQHPQNDYRPRASPSRTHYLGNGOMRTGPMWCSSTGAGSGQRRVYVQDENGPT
ANDCVERIKPDEORACESGPCOMAYGWMGCTKLCGGIIRTLVVCORRNGEFPDL
SCETILDKPDRDROCTHACPHDAAMSTGPSSCSVSCGRGHORVNYCMADGSHLES
DYCKHLAKPHGRKCRGRCPRKWRAGASGCSGRCGVQORHVGOCIGTGHILRETE
CNPYTRPRERCQCGPRCPLTYMRAEBQECTKCTGCBESRYRKVYCVNDNKNENHGAR
CDVSKRPDRESCSLQCEYIWTGEMSEVPSWEL"

```

```

BASE COUNT      1471 a      1112 c      1362 g      1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score = 6 Optimized Score = 12 Significance = -0.38
Residue Identity = 17% Matches = 16 Mismatches = 54
Gaps = 20 Conservative Substitutions = 0
Translation Frame= 1

```

```

QLRHGHQPLRLNAGHPIVSYXMGILFYXTTYYGXTRKRGTKOTQHHLXAPRRERALKACMXHLRQK
10 20 30 40 50 60 70
11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
XAOXROEENOSKKMRKDXGWMXSSITKQRLSNRNGIFCLMXDGOHKRREDPQKDKTFEFLSTVCSLGGCR
80 90 100 110 120 130 140 150 160 170 180 190 200 210
-----RLMCN-NVNGVHKRCRTQHTPWADGTECE-----PKGHCKYGFV
11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
QONGFIWKRKSTLYFNPVNCISLYLKRPRYKFNXYCYCELNDOSXXTGWAFHIFGCSNIIKLLPVAAFE
150 160 170 180 190 200 210
BOSRWNPSSXYCCSLNKTGYLSSHOMXYLRPGXTGNHLXKLFLYXXKXIESEYDREPCVYXHASXX
220 230 240 250 260 270 280
QOQKXRRRSXSPACHSGSNFELHQPRLDVQKVXSKIYHVEFRHMLMVFAXPTXIOITLPACPATARHHPLORE
290 300 310 320 330 340 350 360
XTMXIDFTREFSGVPIYDAVOTALVQXROWSTORLPDSAPHTLGHWDGVRAWKALQWILCSQNRGCPDRDWT
370 380 390 400 410 420 430
LGKLESIMNLLQNMWRGHONSHSVQOTRTKKMKILCFTXNEXLIXYLQHGAMSOAEARLPRTXIVCSLXREAF
440 450 460 470 480 490 500
XHOXSASQCALGPXYIQWNSSDEGPVQVYVLOSREHSLLSASRQSDRWNSLWPGHKXLYLCPGLPDASWNRSCFK
510 520 530 540 550 560 570
LKSPEPRKMMGLMWRXFFPMQNSGRNITXSTLMLQYCGPNSSSWCYQXCAAOFLGNRRXOLLSFIRKQXRIIL
580 590 600 610 620 630 640

```

```

AKMKLCCNGKGNSHWECGGRVQWVBCRRKNXNLNRSHARTFASFGVKGAVVORCTLLFEQYSNXXRTSA
650 660 670 680 690 700 710 720
VLLQSNAMASMQXTLDRGFETTKLHQIGIXSAVCFXSKRSASAARMTHTYXTLWYRLXPVAVCCQOEXMKCP
730 740 750 760 770 780 790
VWLGILPHIGHLICQIXQAGWEDXGXXMFLQOPSOQKPKMLRGXHGMLALFCLDXMFKKLXRMDDPEEK
800 810 820 830 840 850 860
YLCQYPRKXCTGXQOMHNSRESYHSEYQVPLSTVYEIRLVRLVGLHLMKRAAXAPGLVSVXRSIKXXNVPX
870 880 890 900 910 920 930
DOANITYDLAAGMCILAGSLGTVOCHLWTLGIPAKSEEMHMDLYSGRXOYLXCSNXTNXXYPLGLXITIMS
940 950 960 970 980 990 1000
SSPSPGNEEKHICQTKNPVAILWLDPLMSHLMRYPDEINQOLPRXEMLCGXRECLCYPAXTSGKRMFCDT
1010 1020 1030 1040 1050 1060 1070 1080
LWMEGLLELLCDLWARKXGNPNSDYCQLOXPRDRSEXYXPGILXPNNXQGLFHVTPSKDPQWLSAPLP
1090 1100 1110 1120 1130 1140 1150
KXGLSPERQOPPHPCARMKPEVNMPLGSMFQYLQWRIPACCCMSGXKWT
1160 1170 1180 1190 1200
6. US-09-972-467-2 (509-578)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pbLuescriptII SK plus
clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hiroawa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
2 (bases 1 to 5139)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JUN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
source
CDS
gene
FEATURES
location/qualifiers
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/tissue_type="brain"
/clone_lib="pbLuescriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="KIAA1312"
/note="Start codon is not identified."
/codon_start=2
/product="KIAA1312 protein"
/protein_id="BA92550.1"
/db_xref="GI:7242979"
/translation="NSEHTAVISLCSGMLGTFRSHDGYETIEPLQSMDEDEEONK
PHIYRSAPQREPSTGRHACDTSSEHKRNRKDKKTRARRKMGRIINAGDVALLNG
LATEFASVGNKTDNTRKTRHRRKTRFSLPREVEVLVADNMYVHGHNLOHTL
TMSIVASITYKDPSTGNINIVINLVLIHNEODPSTSEFAOTTLLKPNOMOSKNS
PGGHHIDTAVLLTRKODICRAHDCDTGLABELGTICDPYRSCSISESGSLATTIH
ELGAFENRPHDNNKCKEEGVKSPQHWAPLNTTYNEMWMSKSRKYITIEFLDTG
ECLINPEPSRPLPYVLPGLILVNNKCELIFFGQVQCYMMCKRMLNNVNHG
KGCRTQHPADGTECEBPCKHCKYGFVCKEMADYVLDGWSGSPGTSRCSGGI
KTAIRECNRPEPNKGGKCYGVRMKFKSCNTEPKDQDRDRDSCAHFQKHNING
LLPNRWPKYSGILMDRCKLFCRVAGNTAYOLDRIDGTCGGDINDICVQGLC
ROAGCDHYLNKARDKCGCGGSDNSCKYVAGTGFNTYHYVNTVTRPGATIDVR
OHSFEGEDDDVTLLSSKGEFLINGVYTMAREIRINAYEVSSESTAYERN
STRIEQLQLQVLSVGLYNPDVATSNITIEDPQOFTNNSGPMWACKPQOGER
KRKLVCRESQDLVSDQCDRLPQPGHITPECCDILRNHVASRSECAQCLGTR
TLDIYCAKYSRLDKTEKVDGFCSSHPRKPNREKSGECNTGWMRYSAWTECSKD
GOTORRAICVNTRNDVLDSDSKCTHOEKVITIQRCSEFPCQMKSGDSECLVTCGH
KHROVCOFGEEDRLDMCDPEETKPTSMOTCOOPFCASMOAGPWGSCVTCGGYOLR
AYKCIITGYMSYVDNDCNATRPDTODCELPSGHPPAPAPETRRSTYSAPRTOMR
GSMTPCSATCGKCTRMRYVSCRDENGSAVDSACTTLPRAKEECSVTPCGOMKALD
WSSCSVTCGGQKATQVACVNTSDHVIDRSECDQDYTEIDQDSMSFCPORTPSGL
AQHPQNDYRPRASPSRTHYLGNGOMRTGPMWCSSTGAGSGQRRVYVQDENGPT
ANDCVERIKPDEORACESGPCOMAYGWMGCTKLCGGIIRTLVVCORRNGEFPDL
SCETILDKPDRDROCTHACPHDAAMSTGPSSCSVSCGRGHORVNYCMADGSHLES
DYCKHLAKPHGRKCRGRCPRKWRAGASGCSGRCGVQORHVGOCIGTGHILRETE
CNPYTRPRERCQCGPRCPLTYMRAEBQECTKCTGCBESRYRKVYCVNDNKNENHGAR
CDVSKRPDRESCSLQCEYIWTGEMSEVPSWEL"

```

LADPAASAKTQDREKTRKTRKRLSLRPREFVLYVADNRVSTHGENLOHYIL
 TLMSTVASIYKDPSTIGNLNIYIVNLIVIHNEODPSISEFNAQTILKMFQWOSKNS
 PGGIHHDTAVILLTRDQICAHDKCDTGLAELGTICDEPRSCSIDENGLSTAFTHAH
 ECLHFNEMPHDNNKCKEKGKSPQHVAMPNTFTNPMWMSCKRKYITTELDYGV
 BQILNEBESRYPPLPYOLPGLIYNNKOCELLFGPSOVCPMMOCHRLMNNVGVH
 KCRTOHPWADGTCEPEKHCYKGFVCPKENDVYDTGSMGNSMSPCTSCSTGGGI
 KTAIRECNREPKNGKGYVGRMRKFKSCNTEPCIKOKRDEDECAHFDGHEFNG
 LLPNRWPKYISGILLMKDKCLCEKRVAGNTAVYQLRDVTDGTGQDPTNDCVGLC
 ROAGDCHVANSKARDKCGVCGDNSCKVTAGTNTVHAGNTVRIIPAGATINDVR
 OHSFGEEDDDNYLALSSCKGFEFLNGFNVMARELRIGNAVVEYSAPVARIN
 SYDRIEOLLELLQVLSVKLYNPDRVSEFNIJEDKQOOFYNNMSHCPMAGCKPOGGER
 KRKYCTRESQDLTVSDORCDRLPQGHTEPCGDDCLRMHVASRSCSACCGIGYR
 TLDIYCAKSRDLGCTEKYVDFGCSNHPKPSREKSCSECTNGRISAMTECSKCD
 GGTORRAICVTRNDVLDSDCKTHQEKVITIORSEFPCPMKSGDMSCELVTCGKH
 KHRVOYCOGFEDRLNDRMCDPEKPTSMQTCQOPECAWQAPMOCSEVTCGQOLR
 AVKCIIGTMSVVDNDNCNAARPTDQCELPSCHPAPARESTRYSAPRTOMF
 GSWTPCSATCGKGTGRVYSCDENGSVADEACVTLRPRPAKEGSCVTPCGOMKALD
 MSSCVTCGQGRATROVMVNSDHIYDRSEDOOYITETPODCMSKCPORTPSGL
 AOHPRONEDYRPRRSASPARTHYLIGNOMKRTGPMWCSSTCAGSGSRVRYVOCDEGYT
 ANDCYERIKPRDQRACSESGPCOMAYGMWBECKTLGSGGICTRLRLVYCGRSNGERPD
 SCETILDKPRDQCNTHACPHDAWSTGPMWCSGSGRHKORVRYCMADGSHLES
 DYCKHLAPHGHRCRGRCPKMKAGAMSGSVSCSGRQVQORHVGCOQIGTHKIRARETE
 CNPYTRPESREDCQGRPLYTWRAMEOECTKTCGEGSRYRKVVCVDNKNNEVHGAR
 CDVSRPVDRESCSLQPCPEYVWITGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
 Initial Score = 5 Optimized Score = 11 Significance = -0.42
 Residue Identity = 18% Matches = 16 Mismatches = 50
 Gaps = 21 Conservative Substitutions = 0
 Translation Frame= 3

TPSTRSSASAOCCMAHSGLMGIIILNHYSLMMNKKMRNKTNPSTFTGAAPRPSPOEGMHVTPONTKI
 10 20 30 40 50 60 70
 GTVTRRRKRPQENGKGLTWLVYXQHTAXAOORHFLWVIRRTQERKGPREGQVNFPIHGLKXSMWQOT
 80 90 100 110 120 130 140
 TEMFHMEKTFNTIFYLXCOLXLSIKTOVLEIXLILXLTIXLIMNMGLEPYLMLRQHXKTFASGISRR
 150 160 170 180 190 200 210
 TVOVESIMILLFSXQORISAEITNTNIPYAMLNWEPEFVIPIEAVILLVKIVDXVOLLRSPMSNAMCITCLMMT
 220 230 240 250 260 270 280
 TTVVKKKELRVPMSMLQHXSTPTPTGCGQSVVENISLSFXTLVNASVCLTNLNDPTLCLSNQASFTTXI
 290 300 310 320 330 340 350 360
 NNVNKFLDOVLRCAHIXCSADSGAITSMEYTKAAGLSTHPGPMGRSASLESTASMDVFVFKMMSPXOMDP
 370 380 390 400 410 420 430
 GEVGVPLEPAPBEHVEGASQOPESATDQONKQVENVYXVEXNLSPATRSHVSSRSETSEANSVLTLLGSLIL
 440 450 460 470 480 490 500
 TSTVCGPMACAGSLNIVEFYXXRTGASCSAEMOGTPTISFETEXMMLLVARTOMISVSRFAFKLDAIMFYX
 510 520 530 540 550 560 570
 OKFGEINVEFVVAIILIAKQWQEHLLQYIMVTIIMSEFQVLVPIILMCGSTVSOGKOTMTTXYLQAVKVNSC
 580 590 600 610 620 630 640
 X
 QLP-----GIL-----YNNKQCELLIFGPGSOVCPYMMOCRLWC-----NNNVGHHKGCRTG
 10 20 30 40
 XMETLLSQWPKGFALGMLMXSTVGRPLPXKELTQOIALSKNFCRFRCWESCITPMVAIILSIPOKLTINLS
 650 660 670 680 690 700 710 720
 50 60
 HTPWADGTCEPEKHCYKGFV X
 FT-----GTVMGSHGKHAVNPAKNGNENILFAPGNLISLLFLIKDAITGCPSLDTLLNPNVQVQTVTXGGMPLPAGVN
 730 740 750 760 770 780

VVPASAVTAHMTSTVPNINAGMGRILRLMVFAPAIIPNOATVKNAGQVTRVAGAILPGLINVOKAVTVGR
 790 800 810 820 830 840 850 860
 GGGLFVSIPENMMYMMTANAHIKRKRLPFRGAGVSLVHSGNLELTGOSAMSFPVEKGISTASGSVLYKIDXIMIC
 870 880 890 900 910 920 930
 VTLRPSOHLCLRVSSRNHVPGRHVRVBDASVSLVDRDTSXEDXNASLGLICOMXMTMTVMQOLDOLIPRTVNY
 940 950 960 970 980 990 1000
 HHVILLPOLRKRGEAHTHQEPBGDLGLGPHNQLVGVVPGKDTSAAMRRAALNITRVPVLPCLDDQORKNV
 1010 1020 1030 1040 1050 1060 1070
 LKHVYNGRPMWTGALALXLPVGVGQPGKXCVSTVTVTXISGVSYTRILISQKLTFTVPCHHALKGPTVAXLS
 1080 1090 1100 1110 1120 1130 1140
 TPKMKRTIYPCAPAPAAAPMGSEVETSGELAPGPHVVPVYLPADPSGLLVYRKMKTPTQTTVREKNLSKEPY
 1150 1160 1170 1180 1190 1200 1210 1220
 NDLVLSGLMATGESALSCVVEAEXBQDMMSVSGPTVNGFOIXAVKFLINLPVSSVTHMLVHTTLHGVLALG
 1230 1240 1250 1260 1270 1280 1290
 ARVLVLVEGINNEMFTAMOKMEALIXKYITVSTWLSHMGTESAEEDAPNOKLALGVSAICPVAEAYSRCGM
 1300 1310 1320 1330 1340 1350 1360
 AYRSEHTXKPERPSATHTTPDBSRNATAKAHGVPSTLGGQNRGNKAPRPAAYAPGTARMCVMWMTTKTRCMGHA
 1370 1380 1390 1400 1410 1420 1430
 VTXASGRWTVKAVVCPNAPSMASGSOENQORYRPGNCHNRQLSHGLRVAGVYEGVQVOCGRLEATRPICHEL
 1440 1450 1460 1470 1480 1490 1500
 CVFXOVLSFPLKLEFLVCAGEXEMVSKSLQVYILXFTLMGYMKEEIIITGAANSXDAVRCCLKTVKXKILX
 1510 1520 1530 1540 1550 1560 1570 1580
 COGYNKERWOLXOSTSTCCVKAESLSIQLMPCRNVYXRYOTXFFKRGKSGLVCEISXLENGXNFWFXNIV
 1590 1600 1610 1620 1630 1640 1650
 RQKQTPFNQMCACALZTSFGRKXKFSIISEVSOYPPPHISCAQCISRPPLANTINOXXS
 1660 1670 1680 1690 1700 1710

IntelliGenetics

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 099972467-2-19-287-ab037733.res made by jdelaval on Thu 20 Jun 102 6:15:26-1

Query sequence being compared: US-09-972-467-2 (19-287)	
Number of sequences searched:	6
Number of scores above cutoff:	6

KRKLVCTRESDDLTVDSDQCDRLPQPGHITTECCGTCDDLRHMYASSECSAQCGLGR
 TTDITCAKTSRLDGTKEVDGFCSSHPKPSNKEKCSGDMGMYASMTESKSCD
 GGTQRRRAICVNTNDVLDSDCTHQBKVITQRCSEPCPOMSGDMSECLVYCGGH
 KHRVWCQGEEDRLNDRCMDPEETKPTSMOCCOPECSAPAGTWCQSGYGLR
 AVKCIIGTVASVVDNDCAATPEPTDTCOCCPECSAPAGTWCQSGYGLR
 GSWTPCSATCGGTBRMYVCSRENGSVADSCATLPVPVAKRECSVPCGOMKAD
 MESSCTGCGGATROMGVNYSNHYIDRSCODDITPEDDOCSMRPORDPDSGL
 AOHPPONEDYRRRSASPSRTNVLGCMQRTGPMGACSSTOAGDSQSRVVCODENGT
 ANDCYERIKPDEORACESGPCQWATGNNWECTKLCGGGIRTLVVCORNGERFPL
 SCEDILKPDREDQCNTHACPHDAAMSTGPMSCSVSCGQKORNYCNAMKDSHLES
 DYCKILAKPHGHKCRGRCPKKAGAMSGCYSVCGRGVQORHVGCOIETHK IARETE
 CMPTPRESEEDCGRCPLTYTWRAEOMEECTICEGESHRYKRVVCVDKMKNEVHGAR
 CQVSKRPVRESLOPCEYVMTGEMSEVSMEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 11 Optimized Score = 47 Significance = -0.34
 Residue Identity = 20% Matches = 63 Mismatches = 189
 Gaps = 58 Conservative Substitutions = 0
 Translation Frame= 4

MTTGLPLKRVXISIVHRIYERKMLEILRLIKIIFFLKSCLLNHNHSGSXKGFVCLQCFKTNFSPH
 10 20 30 40 50 60 70
 FOVRKSHLLVOIFSLSKYKIRSGNARPHFCMAATAGVKEILLRPSHSMCFVYVAISPCCIPGTTISYTXLSSG
 80 90 100 110 120 130 140
 TSOHPMSLOPLXKFLVLANRTSRXITECRREGTYPRLSITSLPICIGILAMESXGLVRKHTVHSGKSGCQHPQ
 150 160 170 180 190 200 210
 VSHRALIPPLPLPLSGHGXADGSSQDSTPHSHVVIQYSGCKLQLSRSGRLITSRACSTFLLSS
 220 230 240 250 260 270 280
 THTTLRYLESPOVLVHSHSSALQVXRGHRGPWOSRSDGLVYGLHVSALICVDPHXPTCCCTCCTPRPOD
 290 300 310 320 330 340 350 360
 TEHXLOAPAFHLGLHPRPHFLPCPGGLARCLQXSLSKWLPSPAMQXTFRCLCPROETENEHLQGVPLNAASG
 370 380 390 400 410 420 430
 QACVLCRSRGGSLRISQKSGNSPLDRXQTTSLVLMPPRHSLVHSPOLFXHAGGPPDSQALCSSGFTILS
 440 450 460 470 480 490 500
 TOSFAVYPFSSXHTTTRMDPPAOVLHAPQPVYLMHPPTSTWYRLGLALRGRRSSPMKGCXAKPLDGLVIXG
 510 520 530 540 550 560 570
 HGMDSQSWSVSGIXSWSHLSRSTWLSLXLTHTICRVALPCPQVTEOELQSKAFHCPOGVTEHSSPATGLGRV
 580 590 600 610 620 630 640
 AGLSATEPFFSSQOLTYLLIRVPPQVAEHGVODPNRHMYILGALVYLLVSGAAGGKXHGDSQSWSVGYLY
 650 660 670 680 690 700 710 720
 AALQSLSTTDIXVPMHFTALSWYPCPQVTLHCPQSPACODAHSGCXOYCIDVLSGSHILSPNRSPPNX
 730 740 750 760 770 780 790
 HQTRCLCPRPQVTKHSDQSPDFHCGGSGNSLHLXMYTFESXCVHLLSSSTSRVLTOTALLMWPSPOLFHS
 800 810 820 830 840 850 860
 VQEXKRPVYLHSEHFSRLGLGWLLOKPSSTFSPSSLLYLAQXMSNVKPKPMWALHSLLLATCULRS
 870 880 890 900 910 920 930
 X
 10 20 30 40 50
 EMGSPDAANVRKRLRHPQVKKLETL--SEYELVSP--IRVNALGEP--FPTN-----VHKRRTRR
 60 70 80 90 100 110 120
 S--INSTDPWPAFASSSSSTSQAHYRLSAGCQGLFNLTANAGTAPLPTVTLTGCTEVPNQTFR-----
 1010 1020 1030 1040 1050 1060
 SGLVNEPFD--KTXSSSCSMRS--VELLSTAVSDPLSTYTAAPMISLAIYV-----TKPFPSKN

230
 140
 287
 269

130 140 150 160 170
 -----YSEEEALFKCFKGYVNTNSEHTAVSLCSGM-----GTPRSBGDFIPLQSMDEQ--
 SPLLDKAKXLLSSVSPKLCRRSILVAPAGITTYLXPKCTVLANPAPVTLHELSPPORHLSLRAFEK
 1070 1080 1090 1100 1110 1120 1130
 180 190 200 210 220 230 240
 ---DEEONKPHIYRRSAPORPSTGRHACDTSSEKHNRSKDKKTKRARKGGERINTL-AGDVAALNSGLA
 TXSHAPCORPWTQISLFSYMQGVPS---TILSRXXXAVPATLQNNLHR-SFIRIPLVLTGORTLGSRL
 1140 1150 1160 1170 1180 1190 1200
 250 260 X
 TEAFSAVGNKTDNTRERKTRRTRK
 1210 1220 1230 X 1240 1250 1260 1270
 MLKCFPSKXAHCSRRKSPFCLRHGCSVLQDLNFIILRPQVPPFPFGSGLHSMVAMPVPLEQVPGQLQ
 1280 1290 1300 1310 1320 1330 1340 1350
 PODSVTGTSTSLQNTQNYLQCFPGSHVPSAQGVXQRPCLTPPLILLHOSRLHCITYGHTXEPKINSH
 CLFTLXMPGCSWTGKXGLDGSLSKSHSPXVSKNSVLYFRLHFDHLOGLVXKFSVGAMTCGLTFPSSLHL
 1360 1370 1380 1390 1400 1410 1420
 LLSXGMLNTPSSMAIVKAVLANPLSLLEQLAGSQMVPESAPRKSHLSALQISCLVAKRTVASKIIPPG
 1430 1440 1450 1460 1470 1480 1490
 LEFECCHWQKFFNVVXALKDMEGFSCLXITIKFTITILLIKFPIIGSLIXPATIDIKVIXCFRSPWYETI
 1500 1510 1520 1530 1540 1550 1560
 LLSATTKSTNRGDKKRFVILWFLSVLSVLLPXAENASVAPRLFNAMSPARLLISPFILAVLFFLSL
 1570 1580 1590 1600 1610 1620 1630
 CLFLCSEVSHACLVEGSLWGLRLXMMXGLFCSSSSSCSIDSGSIKXSPSXDRNVPSPICEORLMTAVCS
 1640 1650 1660 1670 1680 1690 1700 1710
 E
 3. US-09-972-467-2 (19-287)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_1b:phluescriptII SK plus
 clone:fh11767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hiroseawa,M. and Ohara,O.
 TITLE Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 JOURNAL DNA Res. 7 (1), 65-73 (2000)
 MEDLINE 20181126
 REFERENCE 2 (bases 1 to 5139)
 AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 FEATURES
 source 1..5139

AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail: odana@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)

FEATURES

source

1. .5139
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/issue.type="brain"
/clone.lib="pbluescriptII SK plus"
1. .4417
/gene="K1A1312"
/gene="K1A1312"
/note="Start codon is not identified."
/product="K1A1312 protein"
/protein_id="BA92550.1"
/db_xref="GI:7242979"
/translation="NSEHTAVISLCSGMLGTFRSHDGYFIEPLQSMDEDEEENK
PHIYRSAPQREPSGRHACDSEHKNRSHSKKTKTARKKGERINAGVAAALNSG
LAPFAASAYGKNTDMTRKRRHRTKRPISYRPRVYLVADNRVSTHGENLOHTIL
TLASTVASYIKRPSIGNLINIYVLIYTHNDDPSISENQTTLNFCQOHKNS
PGIHDHDAVLTRDIDICAHDKCDTGLAELGTTCDYRSCSISEDGLSTAFPIAH
ELGVFNMPHDNNKCKEKGVSQHVMAPIENFTNPMWMSKSKRYITFELDTGYG
ECILNEBESRYPPLPOLYGLIYNNKOCLEIFGSGVCPYPMQCRILMNNVGVH
KCGRTQHTPMADGTECEPGRKKGCEVYKREMDVITVSGMSMSPFGTSGTGGGI
TLAIRBCNREPKNGKGCYVGRMKFKSGNTEPDKOKRPDDECAHFDKHNIRING
LLPVRWVYKSGILMKDRCKLEFCVAGNTAYTQLRDYIDSTPGQDNTDLCVQGLC
ROAGCDHYLNSKARDKCGVGGDSSCKYVAGTNYHYGINTVIRIPATNTIDVR
QHSFGEITDDVNLALSSKGEELLNGFNVTYMAKREIRIGNAVVEYSGSTAVERIN
STDRIEBELLDVLSVGLYNPDYVSNIPLEDPQOOFYNSHSGPMQACSKPOGER
KRKIVTRESDDLTIVSDORCDRLPQGHITTEPCGTCDDLHMHVARSFCSAOCIGYR
TLDIYCAKYSRLDGTKEKYVDGEGFSHPKPSMRKSGECNCTGSGRYSANMPECSGD
GGTORRAICVNRNDVLDLDSCTQHEKYVTTORCEFPQPMKSGDSECLVTCCKGH
KHROYWCQGEEDRLNDRKMDDEPTKPTSMQTCQAPSCAMQGPWCQSCVTGQGLR
AVKLIIGTYMSVDDNDNCNATRPDTQDCELPSCPPAPAPETRSYVSPRIQWR
GSMPCSATGCGGTFMRVYSGDENGSVADESACATILRPVAKERC SVTPCGOMKALD
WSSCVTCGCGATROVMCVNYSHVITDRSECDODYIPETPODCSMSCPORPTSGI
AOPRONEDYRPRASPSRFTYLGSMQRTGPMWACSTGAGSGSRVYVODEGTYT
ANDCYERIKPRDQACESGRCPOVAYIGMWGECTKLCGGGIRLRLVCCRSNGERPD
SCETIDKPRDQCNTHACPHDAWSTGWSVCYSVSGGRKHQNRVYVMAADGSHLES
DYCKHLAPHGHRKRCGRCPKMKAGAMSOCSVSGRCGVQORHVGCOIGTKRIARETE
CNPYRPSERDCCGRCPLRYTWRAEWMQECTKTCGEGSRYRKVVVCVDDNKNEVHGAR
CDVSKRPVDRSCSLQCEYVWITGESEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
Initial Score = 8 Optimized Score = 42 Significance = -0.41
Residue Identity = 21% Matches = 61 Mismatches = 153
Gaps = 69 Conservative Substitutions = 0
Translation Frame= 5

EMGSPDAAAVRKRDLRHPROVKLLFTLSEYIYSPTRVNALEPPTNVAHFKTRRSINSADDPWPAFASS-
XLLLVVYHKKSR-XAL-----CTGRYMK--KYLIVFGYIRKLLSSSE
10 20 30 40 50 60 70
80 90 100 110 120 130
-SSSTSSQAHYR-LSA-----FGOQFLNLTANAGFIAPLFTVTLTGTPGVNQTKEYSEEEAEK
140 150 160 170 180 190
AVCXTGTTHLVHBRGLFAPYNNVLKPKISAPIFKLNLTYXSRFLASLESGL-----VLLDHISANPQLEKX
50 60 70 80 90 100
HCFYKGVYNNSEHTAVVISLCSGMLGTFRSHDGYFIEPLQSMDEDEEENKPHIYRSAP-----

-----RFGSLHTACSLSQLPSLLVSLALQYF---LILNCIQ-----AFHSILTYCSCPNNEFFMP
110 120 130 140 150 160
200 210 220 230 240 250
---QRE-ESTGRHACDISEHKNR-HSKDKKTKTRARKGE-----RINLAG---DVAALNSGL-ATEAFSA
170 180 190 200 210 220 230
IGHQGSQNVDLKGLDLHLHLSRFAVKEFLGKAEDLSSENTQFTVANRAGSLKSPFLHSHSLCHSQA
260
YGNKTDNTRKERTHRTKR X
MAELLTWT-VPRVTPLTILXSRHTRRVANVSFHCPRAPACSHSVPAHPRCCPHRTPPCGTWSLRKSMCI
240 250 X 260 270 280 290 300
LAIPDRSKCRDPTVGIGSRVPTPVWCMQGTSLMYLCFVFRSDSPRASAVALRIGHRQSTDSKROLSIMIFLL
310 320 330 340 350 360 370
GFFCAHVAHPGAYSNHFLNGFHLPLCSKHFFVVALDHRKQNTSSKQYSMQRYVDKHVCTAHDRVYQEPH
380 390 400 410 420 430 440
SSNLEFVHMTADRPVLELCIHTTAXCTLPSCHKPTEDKGRILHRLFAHOYLFSPHSLRGIHHPDIOHA
450 460 470 480 490 500 510 520
AGTRQHRVYNNMLPRQGFSTGFHRAHCGSGWRSQDPSFGRGAELSHCLCSFEGMTWNSPQGLGYNPGH
530 540 550 560 570 580 590
THSDRSRCHSGXHTSLAGLPYLAHRSQKSSSPRSIHAHVSONILPRLVXAGXHRHSRQPOSHLGSXR
600 610 620 630 640 650 660
ISSGVLSHKSLMSKSTQIANGFLVHCMSFSGQLCEDDMYTHSPGYQLXHLNHSCHLPLTYKSOXC
670 680 690 700 710 720 730
ISLLAGLIPVHKXCTVPRDPRVARNHPIRADKSAKMLAMSOGHPTFYLIDLHODTRGGAVALFHRXPSPL
740 750 760 770 780 790 800
TSLQISTYDKGTGHTCSXKXLSLDVCICCHPVNHHFGYHKKRPFSSGSHRSFSLIQSRONSAPHCYIPLSTF
810 820 830 840 850 860 870 880
HCLVWDGCKKHHQPSOSSHPACIYHMSRCPMCGNPSHTGHTHYHSCMONHTSGHSLYHRQXCVQAGADR
890 900 910 920 930 940 950
IFDQKQADQIOWCKQVFEVSVPLGAVYOMLAMAHDSSKTAEVYLQLEFYKXNSVNRGCTTFRPRKAYULA
960 970 980 990 1000 1010 1020
QCDLSXFFELROSFRNCHTLPOHSGCEFPFWPLXOOSFHLARLHLYCLIKLSSGRFLPRRCANAHQYWXHQ
1030 1040 1050 1060 1070 1080 1090
LEFGQYCNHNVLYXMLEPLFCQMKNYRHHKPHIYLSGLLSLKHDRIDLAGKPGHRYNLCPGKREPHLSLCL
1100 1110 1120 1130 1140 1150 1160
EADSRILSLPLCRCTCTGSPSPFHCIXGPSAHWEADRCXCNASQSEHTVYLGSLASXIDMAPCCRTYISFY
1170 1180 1190 1200 1210 1220 1230 1240
VLHSTFHHPLV
1250

5. US-09-972-467-2 (19-287)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for K1A1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone.lib:pbluescriptII SK plus
clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (sites)
JOURNAL       Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.
MEDLINE       Prediction of the coding sequences of unidentified human genes.
AUTHORS       XVI. The complete sequences of 150 new cDNA clones from brain which
JOURNAL       code for large proteins in vitro
MEDLINE       DNA Res. 7 (1), 65-73 (2000)
TITLE         2 (bases 1 to 5139)
AUTHORS       Ohara,O., Nagase,T. and Kikuno,R.
JOURNAL       Direct Submission
TITLE         Submitted (31-Jan-2000) Osamu Ohara, Kazusa DNA Research Institute,
JOURNAL       Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
JOURNAL       292-0812, Japan (E-mail:cdna1fo@kazusa.or.jp,
JOURNAL       URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
JOURNAL       Fax:+81-438-52-3914)
FEATURES      Location/Qualifiers
source        1. 5139
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="fh11767"
              /issue_type="brain"
              /clone_lib="pbluescriptII sk plus"
              /gene="KIAA1312"
              /gene="KIAA1312"
              /gene="KIAA1312"
              /note="Start codon is not identified."
              /codon_start=2
              /product="KIAA1312 protein"
              /protein_id="BA02550.1"
              /db_xref="GI:7242979"
              /translation="NSETHTAVISLSCGMIGTFRSHDGYFIEPLQSMDEQDEEQRNK
              PHILYRSAPQREPTSGRHACDTSSEHNRSKDKKTRAKMERINLAGDVAALNSG
              LATFAFAYGKNDTERTKTRKFLSYPPFVVLVAVNRMVSYGENLQHYIL
              TLMSTVSIYKDPISGILNIVINLIJHNEQDGSISFNADTTLKNCOMOHNSN
              PGGHIDTAVLTTRQDICTRAHDKDPTGLAELGCTIDPYRSCSISDSDLSVAFTAH
              ELGHVFMHPHDNNKCKEBCVSKPOHYMAPTLAFYTNPMWMSKCSKITYTERLDYCYG
              ECLINPEPSRYPLVQDGLITVNRKOCILITGPDSQVCPTMQRRLMKNVNVVH
              KGCRTQHPMADGTCEPGRHCKYGCVPREMDVPTDTSWGSWSPFGCSRTGGGI
              KTAIRECNREPRKNGCYVGVRMRKFKSCTEBCLQKRFREDQCAHGDGHFNING
              LLPVNRWPKYSGLIMKDRCLFCRVAGNATYVQLDRVYDGTVPAGDNDIVR
              ROAGCDHVLNSKARDKGCGVGGGNSCKTAVGTFPTVHGVYTVYVTRIPAGNIDVR
              OHSFSGEDDNYLTALSSKGEPLINGNFVMAKREIRIGNAVVEVYSGSETPVAEIN
              STDRIEDELLOVLSVGRKLYNPDRYSFNIPEDKFDQFTLWNSHGFWQCSKPCQGR
              KRKLVTRESQDLVSDQRDLRPLQPGHITTEPGTDCLRWNVASSECQAQGLSYR
              TLDIVCAKYSRLDGTEKVDGFCSSHPKPSNRKCSGECNTGMYRSMATBESKSD
              GGTORRAICVNTNNDVLDSDSKTHQEKVTIQRCSEFPCPOMKSGDMSBCLVTGCGH
              KHRQWCOGFEDRLNDRMCDPEETKPTMOTCOPEASMOAGMGQCSYTCGGVYLR
              AKVCIIGTVMSVVDNDCNATREPTDODELSCHPAPAPETRRSTYSAPTORFE
              GSNTPCSATGCKGRMYVSGRDENSVAAESACILPRVARECSYPCQOMKAD
              WSSCVTCGGGRATROYVCVNSDHYDRSECDYDIPETDDCSNPPORTPDSGL
              AOHFQONDIYRPSASPSRTHVLGNGWRTPGWACSTGASORRVVVCQDQENSTY
              ANDCVERIKPDEORACEGCPQMAVNGMECKRLCOGGLRTLVLYCOSNBEREDL
              SCETIDKPPREDQCNTHACHPDAASTGMSVSGSRGCHKRNRYCAKGSHTLES
              DYCRLAKPHGRKCGRCRCKMKAAGMSQCSVSGCGRVQOHRVGCQIDGTHLARTTE
              CNPYTRPESRDGCGPRCLYTNRAEMOECTCTCGCGSRVRYKVCVVDNKNKNEVGAR
              CDVSKRPVDRSCSLQPCFVWITGEMSEVPSNML"
BASE COUNT   1471 a 1112 c 1362 g 1194 t
ORIGIN
AB037733     length: 5139 June 20, 2002 06:02 Type: N Check: 4803
Initial Score = 7 Optimized Score = 42 Significance = -0.43
Residue Identity = 19% Matches = 66 Mismatches = 149
Gaps = 121 Conservative Substitutions = 0
Translation Frame= 1
              X 10 20
              EMGSPDAAAVRK-DRLHPRGVK

```

WKPRKXLLXAPGXATW
1320 1330

> O <
0110 Intelligenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 09972467-2-289-478-ab037733.res made by jdelaval on Thu Jun 102 6:16:17

Query sequence being compared: US-09-972-467-2 (289-478)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of US-09-972-467-2 (289-478) with:
File : ab037733.seq

```

100-
N -
U -
- -
- -
- -
R -
O -
F 10-
S -
E 5- *
U -
N -
C -
E -
S -
- -
SCORE 0 21 42 63 84 106 127 148 169 190
STDEV 1 1 1 1 1 1 1 1 1 1

```

PARAMETERS

```

Similarity Matrix      Unitary      K-tuple      2
Translation Frame      1          Joining penalty 20
Mismatch penalty      1.00        Window size    32
Gap size penalty      0.05
Off score              0
Randomization group    0

```

SEARCH STATISTICS

```

Scores:      Mean      Median      Standard Deviation
            37          9          74.55
Times:      CPU
            00:00:00.00      Total Elapsed
                        00:00:00.00
Number of residues:      10271
Number of sequences searched: 6
Number of scores above cutoff: 6

```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Init. Opt.

Sequence Name Description Length Score Score Sig. Frame

1. ab037733 TOIG of: ab037733 check: 480 1712 190 190 2.05 2

The list of other best scores is:

Sequence Name	Description	Length	Score	Score	Sig.	Frame
2. ab037733	TOIG of: ab037733 check: 480 1712 8 30 -0.39 4					
3. ab037733	TOIG of: ab037733 check: 480 1712 8 34 -0.39 3					
4. ab037733	TOIG of: ab037733 check: 480 1713 8 28 -0.39 1					
5. ab037733	TOIG of: ab037733 check: 480 1711 7 37 -0.40 6					
6. ab037733	TOIG of: ab037733 check: 480 1711 6 33 -0.42 5					

1. US-09-972-467-2 (289-478)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pluescriptII SK plus
clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 7 (1), 65-73 (2000)
MEDLINE 20181126
REFERENCE 2 (bases 1 to 5139)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@cdna.kazusa.or.jp).
URL: http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

FEATURES

source Location/Qualifiers
1..5139

/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="fh11767"
/tissue-type="brain"

/clone_11b="pluescriptII SK plus"
1..4417

/gene="KIAA1312"
c1..4417

/gene="KIAA1312"
/note="Start codon is not identified."

/codon_start=2
/product="KIAA1312 protein"

/protein_id="BA92550.1"
/db_xref="GI:7242979"

/translation="NSEHTAVISCSMCTFSHGQDFTIEPLSGMDEDEDEBQNK
PHILYRSPQREPSTGRHACDYSSEHNRSKSKKTRAKWGERINTLADVAALNSG
LATPFAISYGNKTDNTRERKTHRTKRLSPREVLVVAADRMVSYGENLAHLYIL
TMSIVASISYKDPISGNILINIVNLIVINEDQDPSISFNAQTILKFCOMQHSKNS
PGGIHDTAVLITRODICRAHDKDGLGLAETICDIPYKSCSISDSGLSTAFTTAH
ELGIVFNMPHDNNKCKKEGVKSPQHMATLNTNPMWMSCSKRYTTEPLDYG
ECLNPEPSRYPLPYDLPGLIVNQCCLILGRPSQVCPYAMQCRRLACNVNVGH
KGRQHTPMADGETCEPGRKCKGFCVPEKMDVPYTDGWSMSSTFGTCSRTCGGCI
KTAIRECRNPEPKNGKRYCVGRNRKFEKSCTEPLQKQDFRDECAHFDGKHFNING
LLPNVRWPKYSGLIMKDRCKLFCRVAGNTAVYQLDRVIDGTGPGCDNDINDICVGLC
RQAGCDHVLNKAARRDKCGVGGDNSSCTVAGTFMTVHYGNTVVRIPAGATNIDVR

ELGHVNMHPDDNNCKEBCYKSPQHVAAPLTINFTNPMMSCKSRKITTLEFLDTGYC
 ECLINPEREPRLPEVOLDGILYNNKOCCELLFGPGSOVCPYMMOCCRRIMCWNVGVH
 KCCRTDHTPMADTCEBEPGRHCKYGEVCEKEMDVYTDGSMGSHGPFSGTSCKPGG
 I KTAIRCNBPENKNGKCYCGRRMKFSCNTPCKOKRDEDEGCAHHDGHEFING
 LTPNVAWPKYSGIIMDKRCLFCRAGNATAYOULDRYIDDTPGODNDCTCYGLC
 ROAGCDHVNASKARDKCGVGGNSCKTVGTNTVYHGYNTVYRIIPACATNTDVR
 OHSFGEEDDDNVIALSSKSGPELLNGNVTYMARREIRIGNAVVEYSGSETAVRIN
 SVDRIEDELLOLVLSGKLYNPVYRSNITPLEDEPQGVNHSHPWQACSPCGGR
 KRLVCTRESDDLTIVSDORDLPQGHITTEPCGTDGDRHMHVARSBSQACGGYR
 TLDIYCAKTSRLDGTEKVDGSCSHSPNREKSGECNTGECWASAMTECKSKCD
 GGOTRRARACVETNRNDVLDSDSKTQHEKVTIQRCESEFPQPMKSGDMEECLVTCGKH
 KHRQWVCOGEDRLRDMCDPEKPTSMQTCOQPCASMOAGPMGCGVTCGOGYLR
 AVKCIIGTWSYVDDNDCNAATRPNTDDBCELPCHPPRAAEPFRSRYNSAPRTOMR
 GSWTSCATCGKGTBRVYSCDENGSVADBSACTILPRPAKECSVTPCGOMALD
 NSSCSVTGCGGATROVMCVNISHVIRSBEDDYEITPDQDCSMSPCPTDPSGL
 AQHPONEDYRFRSPSPRTHYLGNOMTGTAGSCSTAGSGSORRVAVOCODENGYT
 ANDCVERIKPDBORACESGPCOMAYGNWGETKLCGGGIRTLRVCCORSGNERPDL
 SCEIILDKPDRCCNTNACPHDAWSTGPMSSVSCGRCGRHORVYCMADGSHLES
 DYCKHLAKPHGRKCRGRCRCPKAKAGMSOCVSCRGVGOORHVGCOIGTKIKIARETE
 CNPVRPESERDCGGRCPITYMRAEMOECTKCTGCBESRIRKRYVCVDNNNEVGAR
 CDVSKRPVDRSCSLOPCEYVITTEKSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
 Initial Score = 8 Optimized Score = 30 Significance = -0.39
 Residue Identity = 19% Matches = 45 Mismatches = 112
 Gaps = 79 Conservative Substitutions = 0
 Translation Frame= 4

MITTGLPLKRYXISIVHRKIVKMLEILRLIKTIIFLXSCLLNRHNTSGSKGFCVPCLOCFKTNFSPH
 10 20 30 40 50 60 70
 FOYRKSHIIVQIFSLSKIRSGMARPHFCATAGVKEILLPSHSCXEVTVVALSPCCIFGTTIFSTYXLSG
 80 90 100 110 120 130 140
 TSOHPMSLQPLXXFLLYANMTSRXITECREPTGRPSLTLPLCIGILAMESXGLVRKHTVHSGSGXPO
 150 160 170 180 190 200 210
 VSHIPALIPLEPLPLPSGHGXADDGVSODGTSDDHSPVLOTYSQCGKLOLDSSTGGLLISQARPCISFLLSS
 220 230 240 250 260 270 280
 THTTLLRYLESPQVNLVHSGHSSALQVYRGHRGFWSRSDSGLVYGLHVSALAILCVPIXOPTCLCCTPRPOD
 290 300 310 320 330 340 350 360
 TEHXLOAPAFHGLHLPBRHFLCPCGLARCLQXSLSKMLPSFAMQXTFRCLCPROTEHELOGPVLHAASCG
 370 380 390 400 410 420
 10 20 30 40 50 60
 EVLYVADNR-----WVSYHGENL---OHYITLIMSTIVASTYKDPSTIGMLINIVYLVLIHNEODGP-
 70 80 90 100 110 120
 OACVLCSSGSLGSLRISOLKSGNRSPLDLXQTTSLVLMPPHSLVHSP-----QLPXAHXGQ-GPD
 130 140 150 160 170 180
 SOALCSSGTLST-----OSFAVYPPSSXHTTTRRMDPPAOVLEHAPOGVLHWPFPSTWVRLGALRG
 190 200 210 220 230 240
 500 510 520 530 540 550
 -----RKSQ--FMWGC-----XAKPLSGVLAGHDMEQSMVSGISXMSHSLRSITWSLXLTHTICVVALP
 560 570 580 590 600 610
 180 X
 LNFYTNPMWMSKC
 ||

CPOVTEQELQSKAFHCPQGVTEHSSPATGLGRVAQALSSATEPSSRQLTLYLIRVFPQVAEHGVODENRHM
 620 630 640 650 660 670 680
 VLGALYVILLVSGAAGGCGHDSNOSQSWSVGLVALQSLSTSTDTLXVPMHFTALSWPCPOVTLHCPQGRA
 690 700 710 720 730 740 750
 CODAHSGCXYOVIDVGLVSGSHILSEFNSSPNXHOVTRCLOCFPPOVTKHSDOSPDEHCGGNSLILXKWTPRS
 760 770 780 790 800 810 820 830
 XCVHLLSSTSRVLTOTALLLMWPSPOLFENHVOAEXRQPPVLSHPENFRLCLGWLCKRPSSTSVPS
 840 850 860 870 880 890 900
 SILLYAOWXSNVRKXPHMAHSLILLATCRLRSQSVPOGQSVYCMGWSRSHLXSEVYXSDSLVQTSFRRS
 910 920 930 940 950 960 970
 PMOGLILHCHGFXLPXOMXCGLSIGLKEKRTSGLYNFPDTKTSKSSCSMRSEYELISTVAVSDPLYSTTA
 980 990 1000 1010 1020 1030 1040
 FPMRISLAIYVTKRPFESKNSFLLLDKAKXLSSSVSPBKLCRCRISILVAPAGINTVYLXPCTVLNVPATVL
 1050 1060 1070 1080 1090 1100 1110
 HEELSPPQPHLSLRAFEFKTSHHPACRQRPWTQISFVSWFQGVPSITLSRXXXAVPAPTLQNNLHRSFIR
 1120 1130 1140 1150 1160 1170 1180 1190
 IFLYIGTQRTGLSRPLMKCPFSKXAHGSSKRSRCLRHGSVLODLNFIAPTOYFPFPFSGSLHSBMAVL
 1200 1210 1220 1230 1240 1250 1260
 MBPPIVLBOVPKGLQLPDPPSVTGTSTSLGTONPYLQCFPGSHVSPASOGVXYROPICPTPLTLHQSRLLHC
 1270 1280 1290 1300 1310 1320 1330
 ILYGTXBPGRKINSCHLFTLXRMPGSGTQKXGGLDSSLSKHSRXPYKSKNSVYFRLHPDHIOGLVXKESV
 1340 1350 1360 1370 1380 1390 1400
 GAMTCWGLTDPSSLHLILLSXGMLNTPWSSWAIYKAVLNPLSLLEQLXGSOXWPPSARPKVSHLSXALQI
 1410 1420 1430 1440 1450 1460 1470
 SCLVTRTAVXKMPPLGFLPEECOHMQKFFNVYXALXKDMGSCSLXITIKFTITTLIKFPIIGLSIXIRATDI
 1480 1490 1500 1510 1520 1530 1540 1550
 KVKIXCXKPSPEWYETILLISATTKTSTNRNGDXDKKREVLMLVLESLVSLVLPXAMASVAKPLFNATSPARL
 1560 1570 1580 1590 1600 1610 1620

ILSPH

3. US-09-972-467-2 (289-478)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:PBuescriptII SK plus
 clone:Phi1767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hiroseawa,M. and Ohara,O.
 TITLE Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro

JOURNAL DNA Res. 7 (1), 65-73 (2000)

MEDLINE 20181126

REFERENCE 2 (bases 1 to 5139)

AUTHORS Ohara,O., Nagase,T. and Kikuno,R.

TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba

292-0812, Japan (E-mail:cdna@foekazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)

FEATURES

source

Location/Qualifiers

1. 5139

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="fh11767"

/tissue_type="brain"

/clone_lib="pbuascriptII SK plus"

1. 4417

/gene="KIAA1312"

<1. 4417

/note="Start codon is not identified."

/codon_start=2

/product="KIAA1312 protein"

/protein_id="BAA92550.1"

/db_xref="GI:7242979"

/translation="NSEHTAVISLCSGMLGTFRSHDGFTEPLQSMDEQDEEENK

PHIIRRSAPQREPTGRHACDTSSEKRNHSDKKTARKMGERINLAGDVAAINSG

LATFAFSYGNKTDNTRKRRRTKRLSYPRFVEVLVADNRHVSFHGENLOHYIL

TLMSIVASIKRPSISGNLNIYIVNLVIHNEQDPSISFNAQTILKRCOMOSKNS

PGCIHHDVAVLLTRDICKAHDKCDTLGLAELGTICDYRSCSISDEGLSTAFIAH

ELGVHFMHNDNNKKEGVKSPQHVAPLNTFTNWMMSKCRKYITFELDTYG

BCLINPEBRRPYPLVQVGLIYNVKNOCCELIFGSGOVCPYMMOCRIMGNVGH

KCGRTQHPMADGTECEPGKCKGYFCVCKEMDVYVGTGSMGSMSPFETGCRGCGI

KTAIRECNRRPKNRGKCYCVRKFKSGNTEPCLOKRDREDECAHFDKHNING

LLPNRWPKYSIGILMDRCKIFCVANETAYOLARDICTTGCGDNTDIOGJLC

ROAGCDHVLNSKARDKCGVCGGDNSSCKTYAGTINHYGNIYVRIIPACATINDVR

QHSFGETDDNNYLALESSGFEILNGENVITMARRELIGNAVEVSGSEFAVERIN

STDRIEOELLLOVLSVGLYNPDVRSFNIPLEDKPOQFVNMHSPMOCAPKPCQGER

KKRLIVRESQDLTVSDQRDLPOPGHITBPCGDCDLRMHVASRSCAQCGIYR

TLDIYCAKRSYLDGTEKVDGFCSSHPKSRKESGECNIGRWYSGAMECSKD

GGTORRAATVCRNDYLDSSCTHDEKVTORGEFPCQPKSGDMSECLVTGCKH

KHROYWCQGEERLNDKMDPRTKPTSMOQCPKASMOAGPMGOCSTVGGCGYLR

AVKCIIGITMSVDDNDCAATRPDTDOCELPSCHPPAPAPETRRSTYSAPRTDMR

GSMTPCSATCGATRMRYVSCDENGSAVDEACATLPRVAKESCSVTPCGOMKAL

WSSCVTCCGGRATROYMCVNSDVIDRSECDQYIPEITDQCSMSPQRTPDGSL

AQHPFONEDYRRRSASPSRTIVLGNOMTGFWMGSCITAGSGORRVVOCDENGYT

ANDCVERIKEDBORACESGCPOMAYGNMGECKTLCGGGIIRTLVVCORNGSERPD

SCETILDKPDRROCHNACPHDAAMSTGPMSSGSCVSGRGOORHVCQGTGKIAETE

DYCKHLAKPHGRKCRGCRCPKMKAGMSQCSVSGRGOORHVCQGTGKIAETE

CNPYTRPESERDCQGRCPLYTWRAEMOECKTCTGEGSRTRKVVAVCDNKNREVGAR

CDVSKRPVRESCSLQPCYEYVITGSEVSPWEL"

BASE COUNT

ORIGIN

AB037733

Length: 5139

June 20, 2002 06:02

Type: N

Check: 4803

Initial Score

Residue Identity

Gaps

Translation Frame

Initial Score

Residue Identity

Gaps

Translation Frame

Initial Score

Residue Identity

Gaps

Translation Frame

Initial Score

Residue Identity

Gaps

Translation Frame

Initial Score

Residue Identity

Gaps

Translation Frame

Initial Score

Residue Identity

Gaps

Translation Frame

Initial Score

Residue Identity

Gaps

Translation Frame

SRNHGRRVSGDSNVLVDRDTSXEQXNASLGLICQMXMTVMQDLDLIPRTVYHHVILPOLPRKGE

950 960 970 980 990 1000 1010

AHTVQEPSCDGLGCPHQAOLYGVKVPKQDTSAAEMRMALMTLRVPLPCIDQOMQRKNVLXHPGNGRPWGA

1020 1030 1040 1050 1060 1070 1080 1090

LALXPGVGVGPCKXCVSTTVTTXSIGSVTRIIISOKLITRTPCHALKGBQVAXLSTPSKMTYIPGAPA

1100 1110 1120 1130 1140 1150 1160

PAAPCVETSGELAPGBHVPVPLADPSGLLVYRMKMDPQTTVWRKNLMSKEFPVNDALVLSGLMAGE

1170 1180 1190 1200 1210 1220 1230

SALSCVEAXEKQDMWVSGPVTNGFQIAKAVFLNLPVSVTHMLVHTTLGCVLALGARVLSLVGGINNE

1240 1250 1260 1270 1280 1290 1300

MFTAAQMKMEATXKVITVSTWLSHMGTESAEEDAPNGKLAGVSGALCPVAEAYSRCGMVAVERSEHTKXPERPS

1310 1320 1330 1340 1350 1360 1370

ATHHPDRRNATAKAHGVPSLTLGGORNGKNAPRAAKAPGTAARMCVMHTTTRCMGHAVTXASGRWTVKAVV

1380 1390 1400 1410 1420 1430 1440 1450

CNPASMSGQENGQRTPRGNCNHRQLSHGLVAVSGVGVQVEGRIEATRPICHCCLVFXOVLSP-KLE

1460 1470 1480 1490 1500 1510 1520

VLVVADNRHVSFHGENLOHYIL--TLMSIVASIKRPSISGNLNIYIVNLVIHNEQDPSISFNAQTILKN

1530 1540 1550 1560 1570 1580 1590

FLVCAQKXEMVKSLLVYILXFTLMS---YMKKEITTAAN---SXDAVRLCKTVKYLIXCGSYNKE

1600 1610 1620 1630 1640 1650 1660

FCQ-WHNSKNSPGGIHNDTAVLFTFRO-DICRAHDKCDTLGLAELGTICDPRYRSCSISDEGLSTAF-TIANE

1670 1680 1690 1700 1710 1720 1730

PMQLKQTSITCCYKGS-----AESLSLQPLPCRNHYXRYGTTFKFKGSGVGLCEISXLENGXNFWPKNYVIG

1740 1750 1760 1770 1780 1790 1800

LGHVF-NMFDHNNKKEGVK--SPQHVMAPTLNFYNNPMWMSKC

1810 1820 1830 1840 1850 1860 1870

KOTPRMNQMCACLTNPSFRGRKXFSIIEVSGVPRF-----HISCAACLSRPLVAINQXSS

1880 1890 1900 1910 1920 1930 1940

1660 1670 1680 1690 1700 1710 1720

4. US-09-972-467-2 (289-478)

ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000

DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.

ACCESSION AB037733

VERSION AB037733.1 GI:7242978

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (sites)

Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M., and Ohara,O.

Prediction of the coding sequences of unidentified human genes.

XVI. The complete sequences of 150 new cDNA clones from brain which

code for large proteins in vitro

DNA Res. 7 (1), 65-73 (2000)

JOURNAL

20181126

2 (bases 1 to 5139)

Ohara,O., Nagase,T. and Kikuno,R.

Direct Submission

Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,

Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba

292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)

FEATURES
source Location/Qualifiers
1. 5139

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/tissue="brain"
/clone_lib="pluescriptII SK plus"
1. 4417
/gene="KIAA1312"
1. 4417
/gene="KIAA1312"
/note="Start codon is not identified."
/codon_start=2
/product="KIAA1312 protein"
/protein_id="BAA92550.1"
/db_xref="GI:7242979"

CDS

/translation="NSEHTAVISLCSGMLGTFRSHDGYFIEPIQSDMEDEDEEONK
PHILYRSAPQREPSTGHACDISEHKNRSHDKKTRARKGERINLAGDVALNSG
LATEAFSAVGNKNTDRTKRTKRLSYRPEVYLVDNRWVSNGENLOHYIL
TMSIVASIKDPSIGNLINVNLIVINEODPSISFNAQTLLKNFCOMOSKNS
PGJHHTAVLITRODICRAHDCDTGLAELGIDCEYRSGSISEDSGLSTAFIAH
ELGHVPMPHDDNNKCKEEGYKSPQHWMAPLNLTINPMMSKCSRKRYTFELDTG
BCLNBPESRYPPLPOLGLILYVKNKCELTIFPGSVCYMMQCRRLKNNYNGVH
KGCRTQHPMADGTECEPEGKCKYFCVPRKMDVAVTDGWSMSPEFTCSRTGGGI
KTAIRECNPPEPKNGKYCVGRMKFKSGNTEPLKORDRDEQAFHDKHFNIG
LLPNRWPKYSGILMDRCKLFCRVAGNTAYOLDRIDGTGQDPTNDICVGLC
ROAGCDHYLNKARDKCGVCGGNSCKTAVGENTHYVYNNVIRIPAGATNIDVR
OHSGFETDDNNTALSSGKGFLLNGFNVTAKREIRIGNAVLEVSGSTAEIRIN
STRIEDELLOVLSVCKLYPDYRISNPIEDKPOQFYNNSHGPMOACSKPOGGER
KRKLCTRESQDLTVSDQRCLPQPGHITPCCGTDCLRNHVSRSFSAQCGLGTR
TLDLYCAKYSRLDKTEKVDGFCSSHKPSNRKCSGECNTGMRYSAMTECSGD
GGTORRAICVYNTNDVLDKCTHOEVTIORSSEPCPMKSGMSECLVTGKGH
KHROYMCOEGEDLNDRCMDDEPTKTSMTQOOCPCASWAGPMWCGCGGOLR
AVKCIIGTVMSVVDNDCNATRPDPDCLPSCHPAPAPETRRSYTSPRTOMR
GSMTCSATCKGKTRMRYVSCRDENGSVADSACTLPRPAKECSVTPGOMKALD
WSGCVTCGGGRATROYWCYSDHVIDRSDCDYLPETDQCSMSPCPTPDSGL
AOHPONEDYPRRSASPRTHVLGNGMRTGPMWACSTGAGSRRVYVODENGYT
ANDCEVERKIPDEORACESGCPOMAYGNMGCTGCGGICRTLLVOCORNGERPD
SCELLIDPRDEOCNTHACPHDAANSTGPMWSGSGRGHRCORVYCMARQSLES
DYCKHLAKPHGHRKCGRCRCKKAGANSQSVSGCRVORRHQCCGTIKIARETE
CNPTTPRESERDCCGPRCPLTWRAEMOECTKTCGGBSRYRKVVYVDDNNNEVGAR
CDVSKRPVDRSCSLQCEYVITGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
Initial Score = 8 Optimized Score = 28 Significance = -0.39
Identical Identity = 17% Matches = 38 Mismatches = 135
Gaps = 43 Conservative Substitutions = 0
Translation Frame= 1

CLMXXDGOHKREDEQKDKTFEFLISVCSLGGGROQNFIPWRKSTIYFNFNWNCSTLYLKRPKYWKNNXY
110 120 130 140 150 160 170 180
CYCELNCDSXXTGMAFHIFXCSDNITKILPVAAFEOSRWNPSSXYCCSLNKTGYLSSROMXYLRPGXTGNH
190 200 210 220 230 240 250
LXSLKLFLPXXKXWIEFSTYDRPXAAGPCYHAASXXQOQXRRRSXESPPCHSGSNTELLHQPLDVAVKXIKY
260 270 280 290 300 310 320
HXVFRHMLRVFAFRTXIQTLPFACPRTARHPLOREXTXIDFWTRRSGVPIYDAVOTALVOXRWSTQRLPD
330 340 350 360 370 380 390
SAHTLGRMDVAKALALOWIILCSORNGCPDRRWILGLESIMNLONMRGHQNSHVSQOQTRFKKWKIL
400 410 420 430 440 450 460
CTXNIEIXVLQIGAMSOAEARLPRTVCSLXREAFYHORSASOQALGPYIIONNSDEGPVOVYVLQSGREHSIL
470 480 490 500 510 520 530 540

SASROSDRPNNSLMPGHKXYLCEGRLPASMMRSCFKLKSPEPKMMGLMMKXFFPMQNSGRNITXSTLMDQCGP
550 560 570 580 590 600 610
NSWCYQYXCAAAOFLRGNRRXOLLSFIKQXRILLAKWKILCCHNGCKGMSHWECCGRVOWMDRCRKNLNR
620 630 640 650 660 670 680
SHXARFFASGPFVGKVVOPRCITLFPQYSXKXRTSANVLLQSNAMASMQXTLPRTGTCTCLHQGTYSACFX
690 700 710 720 730 740 750
SKMRSAAPAMTHYXLTLMVRLXPEVACCOEOMKXCPYWLGLPHIGHLLCOIXOAGMEDXGXWFLQOPSOYK
760 770 780 790 800 810 820
QXKXKLMRGXHGHLALFCUDXMFKKLXKMDPREKGYLCOYPRKXCTGXOOMHRSRESHSEVOXVPLSTYEIM
830 840 850 860 870 880 890 900
RLVRVLGHLMKRAKXAPGLVSVWXSRSIKXXNVXPRDOANIYADLSAAGMCILAGSLGTVOCHLMTGIPAKS
910 920 930 940 950 960 970
SEMHHMDLYVSGRXXOYLXCSNKNXYPGLXITIMSSSPSCPGNEKHIOCTKNPVAIMVLDPMLSHMERYP
980 990 1000 1010 1020 1030 1040
DETROLPRXEMILCGARECLCYPRXTSGKRMCDTILMAEGICLLELLCDLMARXGNPASDVOLOXPRDRS
1050 1060 1070 1080 1090 1100 1110
VIVVADNRWVSNGENLOHYILTMSIVASYKDP---SIGNL---INVLIVINEQ-----DGPS
1120 1130 1140 1150 1160 1170
EXYXPCOLYRNKRG--LPH--VTMPS-----KDPROMLSAPLPKXGSSPPOQPIPCARMKVEWNP
1180 1190 1200 1210 1220 1230 1240
IS-----FNAOTLTKNECQOMHNSKNSPGJINHDTAVLITFRODICRAHDKCDTLGLAELGTICDPYRSCS
1250 1260 1270 1280 1290 1300 1310
GSMFYOLKRIAPACCSMGXKXTHHKRILCGENKITYXASL-----XIRPLSSVGLQOLGYNHVAWVRH
1320 1330 1340 1350 1360 1370 1380
PFRKXLLXAPXATMAQKVPKRRKMPOMESMRLESYLCVLMPTRTAACGISDRNTQNSORDRYOPTHQGVG
1390 1400 1410 1420 1430 1440 1450 1460
TRLRPRETVSPHLEGRGAMRMODLRRKIQVPOGVGCGXQOKRGAMGTILXREDAAGPYKKIXATILRVGLDH
1470 1480 1490 1500 1510 1520 1530
RMVVGSTVLGTVTVSSAMAXEMORDEMRDECRNMGDLRLPARFATVNCVFSDKSSAPPSXNSLYAKRRECK
1540 1550 1560 1570 1580 1590 1600
RMSLSPRSTFCDSYXCPIGIIKKILLOGIOTHRMLXGAXROLISRYKCSARDTPRBDGCKDLAHVXREON
1610 1620 1630 1640 1650 1660 1670
LFHSSCGHAEMMSVTRPDPFSREAKNLDYVRFPNLKMGAELTFEFTLXGANKPLSXTRCVVAVXOTASEEE
1680 1690 1700 1710
NNFLXLPKLYKHYFFIYLPVHNAYLDPFYKXXTSSNH
1720 1730 1740 1750 1760 1770 1780
5. US-09-972-467-2 (289-478)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978

KEYWORDS Homo sapiens brain cDNA to mRNA, clone_lib:pbluescriptII SK plus
SOURCE clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
JOURNAL 20181126
MEDLINE 2 (bases 1 to 5139)
REFERENCE Ohara,O., Nagase,T. and Kikuno,R.
AUTHORS Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yama, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
source
1. 5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/issue="brain"
/clone_lib="pbluescriptII SK plus"
1. 4417
/gene="K1A1312"
1. 4417
/gene="K1A1312"
/note="Start codon is not identified."
/codon_start=-2
/product="K1A1312 protein"
/protein_id="BAA92550.1"
/db_xref="GI:7242979"
/translation="NSHTAVISLCSGMLGTRSHDGYETIEPLOSMDQEDDEEONK
PHIYRSAPQREPSGTHACTDSEKHNHSDKKKTARKKGERINLAGDVAALNSG
LATEASAGKNTDNTREKTRHRTKRLSPREVEVLAVDNRVSYHGELOHYIL
TLMSIASIYKDPISGNLINIYVNIYIHNODGPSISENNQTLTKNCOMHSKNS
PGIHHDTHVLLTRDIDICRAHDKCDTGLAELGTCDYRSGSISESGSLTAFTIAH
ELGAVENPMHDNNCKKEGYSRQHVAPRTINFTNPMMSKCKRKTITERPDIYAG
ECLINEPESRPLPLVQDGLIYVNVKQDELTFGSSVCPIMQCRKLWCNNVGVH
KCGRTQHPWADGETCEPGKCKYGFVCPKENDVATVGSWMSMSPFGTCSGTGGGI
KTAIRCNRPBERNGKQYCVGRMKFKSCNPEPCKORDFRDEGCAFDKHFNING
LIPNVWVYKSYGILMKDRCKLFCRVAGTAYVQLDRIDICTPGGODNTICVQGLC
ROAGCDHVNASKARDKGGVCGGNSCKTVAGCFNTVHYGNTVRIIPACATINDVR
OHSFSGETDDNYIALSSKGEPLLNQNVYMAKREIRIGNAVVEYSGSETAVERIN
STDRIBQELLQVLSVGLYNPDVRISENIPIEDKPPQFYMNSHGPWQACSKPCQGER
KRLIVCTRESQDLTVSDQCDRLPQGHITTEPCGCDLRHNVASRSCSAQCGYR
TLIDICAKSYRLDGTKEVDGFCSHKPSRREKSGECNCTGWRYSAMTECSGCD
GGTORRAICVNTRVNDVLDSSCKTQHEKVTIQRGSEPCPMKSGMSECIYTGCKGH
KHROYWCGEGEDRLNDRMCDPTKPTSMOTQOQPCASQWQACPMQCCSVTGGCQQLR
AVKCIITGMSVYDNDCAATRPITDQCELPESHPPAAPETRRSYSAAPRTQFR
GSWTPCSATCGKGTMRMYVSCRDENGVADESACATYLRPAKRECSYTPCGQWALD
WSSCVTCGGGATROVMKVNSDHIIDRSBDQDIYIETDQDCMSPCPQRTDPSGL
AQPONEDYPRRSASPRTHVLAGNOMRTGWMGACSTAGSGSRRVYVQDDENGYT
ANDCEVERIKIPDORACESGPCOMAYGNMGMBEYTKLGGGIRTRLVYCCRRVSGEPLD
SCETLIDKPPDRCQCTHACPDAASTGPMSCSGYSGRGRKORVYVQAMADQSILES
DYCKHLARPHGRKRCGRGCRPKMKAGANSQCSYSGRVOOQAHVQCGTGRKLAIRETE
CMPYTRPESERDCGCPCLYTMVRAEEMOECTKTCGSESSRYRKVYVCDNKNREVGAR
CDVSKRPVDRSCSLQPCPEYVWITGEMSEVDSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score = 7 Optimized Score = 37 Significance = -0.40
Residue Identity = 23% Matches = 49 Mismatches = 137
Gaps = 25 Conservative Substitutions = 0
Translation Frame= 6

DYWFETIKKGLDKCAQEDIKKNGTXDTSIIENYFLPLKLFVQAOAHWMKGVCLPLTMFXNOKFOPE
10 20 30 40 50 60 70
SSXSEIHSNPDPEXPLKKNQWXRXTPTFLHGSWSERDSAPTOHLYVCHSCHLSLPLMWHNYILYLVFRH
80 90 100 110 120 130 140
IYAS-ITKDP-SIGMLINIVNLIVIHNEODGPSIS--FNAQTT-----LKNFCOMOHKNSPGGIHNDT
100 110 120 130 140 150 160
AVLLTRDIDICRAHDKCDTL-----GLAELGTICDPRRSGSISDSGL-STAFTIAHELCHVYVNM-HDND
150 160 170 180 190 200 210
LVAYPEAPAVIISLCOXDIKVNHRMXTXDLLETISYISPALHTRNSLCKLRTCKQTHS--SOMOI-GRV
200 210 220 230 240 250 260 270 280
ASSLPHSCTHPSTHPSATLRWLSYRWLQFPGRYLXP-FSCDDPDLAQLQTTAFTVNRPLAHVACPMHLVE
220 230 240 250 260 270 280
160 170 180 190
NKCKEEGVKSPQHVMAPRLNFTYNPMWMSKC
VVIITHHLAVPGFAAGLGAFLPLICPPSVEGTWALAVAFRLSGVWVALGSGYFVCSDDLTAHMDLLYAS
290 300 310 X 320 330 340 350
ATGHALRPSASEPFGASSSSALSVPMWLSOVLVITFTXMASIFCHAVNISLFPSTRTDRTRAPRASTPCS
360 370 380 390 400 410 420
VYWTSMCYVTLTLTIGRFINKFTQAIWKPTTVGRLDHDQSCSTASTTOLSAISPVAISPLKTRAGTGSLLIRE
430 440 450 460 470 480 490 500
YSLHTVCGVSTFILLYNNTPIGSASTGTGCSGSPASPLVSTEHMGAAGAGRTIYLILEGVLSXATVWG
510 520 530 540 550 560 570
PLRAXHGTIVLVSFMDIILVNLITPLIDHYVTVVDTHHLDGCTTLPGRHARAPVQGLPELPTGCHTFELCHMS
580 590 600 610 620 630 640
RQSGTGLTVSHRAILISADVSHPTPTPTSGXAMGPRKSPBLSGNCVTGASPRFGSGRMTXWXFTVLGIS
650 660 670 680 690 700 710
WSSCCTIYVVIYHAIISRNDAFHGXLVLSLSTSDPLSPGRLRQCTFRLLTSLHRCMLGLRVHSHIIXSIF
720 730 740 750 760 770 780
TKLTPDLAVLMPFSGDQALXPVSFRPLMTRBELTAPRLNGEFLMLCAFAVIOYIISGIDTNSPSPLGPTVAF
790 800 810 820 830 840 850 860
XTFSFGRIAPATRTVFPFAFTVAMFGMAAKTIINLSLPIQFAIRGTVVQCAVQATIGTFTPAGMNP
870 880 890 900 910 920 930
POVTVCTGFSNVSRILGQPIASLINSKLIRFPGANKSPFPPLAGFTACLPMDTVPVKLLRIFNMNIER
940 950 960 970 980 990 1000
IAYIVVQLSHRONIKQKFLMLNATGXVNSFTGSLGPTVLYHSIFNANFPFGHCNKKYSIXOEFPTTXAXXV
1010 1020 1030 1040 1050 1060 1070
VIVVEFPPEYVLPPLHINISGTSWNSHSHIVTIMYCIKSCSCHFAAXRIATNPTFISPGFYXNMIIASSLPAK
1080 1090 1100 1110 1120 1130 1140
ALDDITICVLAATRSSIYHSVKSLIVGCPCHSAEQLADVLYHQNSTVFRDPAHIGKQTVDMKMLPVKYSTLEFI
1150 1160 1170 1180 1190 1200 1210 1220
SEVSLLEFTWLRVAGLKFHSTSYVFTSIIFPMWSVALSNGCFDAPSTCSGAGSGKPTSPGICGHDIHFFG
1230 1240 1250 1260 1270 1280 1290
NTKSILAVLSRLALRPIGPGCVL
1300 1310
6. US-09-972-467-2 (289-478)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

> 0 <
01 to Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 09972467-2-ab037733.res made by jdelaval on Thu Jun 102 6:12:58-PDT.

Query sequence being compared: US-09-972-467-2 (1-1629)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of US-09-972-467-2 (1-1629) with:
File : ab037733.seq

100-
N -
U 50-
-
-
R -
O -
F 10-
-
S -
E 5*
O -
U -
E -
N -
C -
E -
S 0-
SCORE 0 163 327 490 654 817 981 1144 1308 1471
STDEV 1 1 1 1 1 1 1 1 1 1

PARAMETERS

Similarity matrix Unitary 6 K-tuple 2
Translation frame 6
Mismatch penalty 1.00 Joining penalty 20
Gap penalty 0.05 Window size 32
Gap size penalty 0
Self score 0
Normalization group 0

SEARCH STATISTICS

Scores: Mean 252 Median 9 Standard Deviation 597.19
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 10271
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name Description Init. Opt. Length Score Score Sig. Frame

1. ab037733 TOIG of: 2 standard deviations above mean ***
TOIG of: ab037733 check: 480 1712 1471 1471 2.04 2
***** 0 standard deviation from mean *****

2. ab037733 TOIG of: ab037733 check: 480 1713 10 219 -0.41 1
3. ab037733 TOIG of: ab037733 check: 480 1711 8 228 -0.41 5
4. ab037733 TOIG of: ab037733 check: 480 1712 8 80 -0.41 3
5. ab037733 TOIG of: ab037733 check: 480 1712 8 154 -0.41 4
6. ab037733 TOIG of: ab037733 check: 480 1711 7 205 -0.41 6

1. US-09-972-467-2 (1-1629)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978

KEYWORDS Homo sapiens brain cDNA to mRNA, clone.lib:bluescriptII SK plus
SOURCE Clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE Nagase,T., Kikuno,R., Ishikawa,K.I., Hiraoka,M. and Ohara,O.
AUTHORS XVI. The complete sequences of unidentified human genes.
TITLE Prediction of the coding sequences of 150 new cDNA clones from brain which
code for large proteins in vitro

JOURNAL DNA Res. 7 (1), 65-73 (2000)
MEDLINE 20181126
REFERENCE 2 (bases 1 to 5139)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.

TITLE Direct Submission
JOURNAL Submitted (31-JUN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/bu/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

FEATURES
source Location/Qualifiers
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/tissue_type="brain"
/clone.lib="bluescriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="KIAA1312"
/note="start codon is not identified."
/codon_start=2
/product="KIAA1312 protein"
/protein_id="BA92550.1"
/db_xref="GI:7242979"

/translation="NSEHRAVIVICSGMCTGRSHDGVFTPELOSMDEDEDEQNK
PHIIRRSAPORPSTGRACDTSEHKNHSDKDKTKRARKGERTINLAGDVANLNSG
LATEAFAYGNKTDNTRKTRTRKRLSPREVLVADNRVSYHGEVLAHYIL
TLMSIVASIKDPSIGNLNIIVNLIVLHNDODGSPISFNQITLKNFCQWHSKNS
PGGIHHDVAVLITRODICRAHDCDTLGLAELGTICDPYRSGSISEDSGLSTAFITAH
ELGVHNMPPHDNDNKKCKREGVSKPOHVNAPTLNFTNPMWMSKCSKRTITTELDGCG
ECILNPEPSRPPLPOLGILLNVNKOCELIFGPGSOVCPYMGCRILMNNVGVH
KCGTQHTPPADCTECBPGKHCYGCYVRKENDVPTDSDSGMSBPGTSCGCGGI
KTAIRECNREPKNNGKCYGVRMRKSCNTEPCLOKORDFDECAHDEGHEING
LPLPVRAWPYSGIILMKDRCKLECRVAGVAYQLDRIYDITPGGDTINDICVGLD
ROAGCDHVLNSKARDKCGVCGGDNSSCKTAGFTVYHGYTNPVIRIPAGATINDIVR
OHSEFGETDDNYLALSSKGEFLNGNEVFMARKEIRIGNAVVEGSGEAVVERIN
STDRIBELLQVLVSGKLINPDVRSFNIPTEDKPOQRYMWSHDPWQACSKPCGGR

```

KRLVCTRESQDLTVSDQRDLRPOGHITTEPCGTDCLRMVHASRECSAOCGLAY
TLDTYCAKYSRLDKTERVVDGFCSSHPKSNRKCSCGECNTGMRISATFESKSCD
GGTORRAICVNTNRNDVLDJDSCKTHQEKVTLORSEPCPOKMSGDMSECLVCGKH
KHROWCOFGEDRLNDRCDBETKPTSMOQOOPESASMOAGPQCSVYOGGYOLR
AVKCIITGYMSVVDNDONATREPTDODECELPBPAPARETRSTYSAPRTOMF
GSMPTCSATCKGKTRMYVSCRDENGSVADBSACATLPPRAKRECSVTPCGOMKLD
WSSCSVTGGGRATROYCWVYSHVYIDRSCDDYLPENDOCSSMPCORPDGSL
AOHPQONEDYPRRSASPSRTHVLGNOIRGTPWAGCSTCAGGSORRYVVOBENCT
ANDOVERIKIPDEORACEGSPOMAYGNWCECYKLCGGGLRTLRVYQORNGEFPDL
SCEILDKPPDEOCNTACPHDAAMSTGPPSSCSGGRHOKNRYCAKDGSHLES
DYCHLAKPHGRKRCGRGRCPKKAGAMSCOSVSGRGOORHYGCOIGTHKILARETE
CNPYTRPESERDCCGRCPLTYMRAEMOECKTCGCSGRYRKVYVCDKMKNEVHGAR
CDVSKRPDRSCSLQRCREYTWITGEMSEVSMEL"
BASE COUNT      1471 a      1112 c      1362 g      1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score - 1471 Optimized Score - 1471 Significance - 2.04
Residue Identity - 100% Matches - 1471 Mismatches - 0
Gaps - 0 Conservative Substitutions - 0
Translation Frame- 2
MOFVSWATLTLTLVROLAEMGSPDAAAVRKORLHPROYKLTLETSEYIYSPIRYNALGSEFPNTVHFKRT
RNSINATDPMWPAFASSSSSSSSSOAHYRLSAGQOFLFNLTANAGFIAPLFTVTLLGTPGVNQTFYSEEE
150 160 170 180 190 200 210
AELKHCHYKGYNTVNSHNTAVISLCSGMLGTFRSHDGDYFIEPLQSMDOEDBEONKPHIILYRRAPQREP
NSEHTAVISLCSGMLGTFRSHDGDYFIEPLQSMDOEDBEONKPHIILYRRAPQREP
X 10 20 30 40 50
220 230 240 250 260 270 280
STGHACDSTSEHNHRSHKDKKTRARKWGERINLADVAALNSGLATEAFSAAGKNTDTRERKTRHRTKRF
STGRHACDSTSEHNHRSHKDKKTRARKWGERINLADVAALNSGLATEAFSAAGKNTDTRERKTRHRTKRF
60 70 80 90 100 110 120 130
LSYPRFEVYLVADNRWVSYHGENLOHYILTLMSIVASYIKDPSIGNLINIVNLIVTJHNEODGPSISFNA
LSYPRFEVYLVADNRWVSYHGENLOHYILTLMSIVASYIKDPSIGNLINIVNLIVTJHNEODGPSISFNA
140 150 160 170 180 190 200
370 380 390 400 410 420 430
OQTLKNFCOMQHSKNSPGIHHDTAVLLTRDODICRAHDKCDTLGLAELGTCDPYRSCSISSEDSGLSTAFI
OQTLKNFCOMQHSKNSPGIHHDTAVLLTRDODICRAHDKCDTLGLAELGTCDPYRSCSISSEDSGLSTAFI
210 220 230 240 250 260 270
440 450 460 470 480 490 500
AHELGHVFNPHDDNNKCKEKEGYKSPDHVNAPLTNFTNPMWMSKSKSRKTIITFPLDTGYECLLNRPESPY
AHELGHVFNPHDDNNKCKEKEGYKSPDHVNAPLTNFTNPMWMSKSKSRKTIITFPLDTGYECLLNRPESPY
280 290 300 310 320 330 340
AHELGHVFNPHDDNNKCKEKEGYKSPDHVNAPLTNFTNPMWMSKSKSRKTIITFPLDTGYECLLNRPESPY
AHELGHVFNPHDDNNKCKEKEGYKSPDHVNAPLTNFTNPMWMSKSKSRKTIITFPLDTGYECLLNRPESPY
510 520 530 540 550 560 570
PLPVOLREGILVNVKQCELLFGPSQVCPYIMOCRRIMCNVNGVHNGGCTQOTHPADGTECEPGRKHCKYGF
PLPVOLREGILVNVKQCELLFGPSQVCPYIMOCRRIMCNVNGVHNGGCTQOTHPADGTECEPGRKHCKYGF
350 360 370 380 390 400 410
580 590 600 610 620 630 640
CVPRKEMDVPTVDSGWSGSPGTCSTRICGGGKIKTAIECNRPKNGKCYGGRMKFSCNTEPCKOKORD
CVPRKEMDVPTVDSGWSGSPGTCSTRICGGGKIKTAIECNRPKNGKCYGGRMKFSCNTEPCKOKORD
420 430 440 450 460 470 480 490
650 660 670 680 690 700 710 720
FRDEOCQHEDCKHNNINGNLPRNVKRWPKYSGLIMKDKCKLFCRYAGNTATVYOLDRKRIYDIDTGPCGQDINDICV
FRDEOCQHEDCKHNNINGNLPRNVKRWPKYSGLIMKDKCKLFCRYAGNTATVYOLDRKRIYDIDTGPCGQDINDICV

```

CGAGVQORHVGCGIGTHAKARETECNPTTPRSESDCCGPRCPDLYTMAEEMOECTKTCGSGSRKRVVCCVD
1360 1370 1380 1390 1400 1410 1420
1590 1600 1610 1620 X
DNKNENHAGACDYSKRPVDESCSLDPCFYWTITGEMSEVPSEWL
|||||
DNKNENHAGACDYSKRPVDESCSLDPCFYWTITGEMSEVPSEWL
1430 1440 1450 1460 1470 1480 1490
GYRPPDPLXTVCFEITSPDLSQARIPCMOSGRDVRDGLVYPSGLSHVHLVDLLAXRMYRGCKLIGCCVEP
1500 1510 1520 1530 1540 1550 1560 1570
EDSVXENITVPCIOGGEATVYTNXHMCEGSRISFTPAVAMQKGLALPDLIPQERLKIWTSMXDFLTXXW
1580 1590 1600 1610 1620 1630 1640
GLKFLVAKHCKGOTNFHPPDVLCLFENKOLQRRKILFFYNIRISITISYFLCLMLITLTFNGKRPVVI
1650 1660 1670 1680 1690 1700 1710
2. US-09-972-467-2 (1-1629)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:pbuescriptII SK plus
clone:fh1167.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hiroseawa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res 7 (1), 65-73 (2000)
JOURNAL
MEDLINE
20181126
REFERENCE
2 (bases 1 to 5139)
Ohara,O., Nagase,T. and Kikuno,R.
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@fokazusa.or.jp)
URL: http://www.kazusa.or.jp/huge/; Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
FEATURES
source
1. 5139
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh1167"
/issue_type="brain"
/clone_lib="pbuescriptII SK plus"
1. 4417
/gene="KIAA1312"
<1. 4417
/gene="KIAA1312"
/note="start codon is not identified."
/codon_start=2
/product="KIAA1312 protein"
/protein_id="BAA92550.1"
/db_xref="GI:7242978"
translation="NSEHTAVISLCSGMLGFRSHDGYEIEPLSDMDEDEEEDKONK
PHIYRSAPORPESTGTHACDTSEHNRHSKKKTKRARKMERINILADVALNSG
LATEFSAVQREPTGTHACDTSEHNRHSKKKTKRARKMERINILADVALNSG
TLMSIVASIKYDPSIGNINIVINLVINREDGPESTNAQTILKNEQWHSKNS
PEGIHHNPLAVLKDICTRAHDKCDTGLGLGTCIPYSCSISEDSGISTATIAH
ELGHVENPHDNNKCKEEGYSQIHVAPLNLTYTPMWSKCSKRYITTEFLDTYG
KCLINEPSRYPLEPOLGILVNVNOCETLIGLAEIGTICDPYRSC-----SISDSGLSTAFITIA
KCGRTQHTPMADGTECEGKCKYFCVPEMDVPYDGSWGSWSPGTCSPFCGGGI

KIAA1312 protein
LIPNRVPRYSGILMKDRCKLCRYAGNATVAYQLDRVIDTPEGODNDICVOGLC
ROAGCDVLNLSKARADKCGYGDNDSSCTVAGTEHTVHYGYNTVYRAGATNDIVR
OHSEGEEDDDNDLALSSKGEFLNLNGNEVYMAKREIRIGNAVEYSGSEFAVRIN
STIDIEBELLOLVSVKGLYNNPVYRSPNIDIKDPOGYMWSGMOGCPGGER
KRKIVCRRESLDLVSDROCDRLPOGHITPDEGTPGCDLRHVAVASRESACCGAYR
TLDIYCAKYSRLDQKTEKYVDGSCSHRPSPNREKSGECNNGRYSMTKSCSCD
GGTORRAICVNTNRNDVLDKCTHDEKATYIORSFPCPKSGMSECTVCGKRGH
KHROVWCQFEDRLNDMDPEETKPTSMOCCQDPEGASWQADPGQSVTCGQYOLR
AVKCIIGTMSVVDNDNCNAATPTDODCELPSCRPAPAEIRKSTYSAPRTQMR
GSMTPCATCGKGTMRMYVSCRDENGSVADESACALPAPAEIRKSTYSAPRTQMR
WSSCVTCGQGRATROVMYVSDHYIDRECGODYIPEPDCCSPPCQOMKALD
KQHPFEDYRBRSPASRTHVYSGNOMRTGPMGASSTICAGSGORRYVCCDENGIT
ANDCEVIRKDEQACSGSGCPQOMAYGMECTKICGGGIRIRIYVCOBSNERPDL
SCETLDPDPRDQCNITACPHDANSTGSPSSVSCSGGKHORVYCAKYSRLD
DYCHLAKPHGRKCRGRCRPMKAGMSQCSYSGRGVQGRVAGQDITHTIARETE
CNPTRESEMDCCGPRCPDLYTMAEEMOECTKTCGSGSRKRVVCCVDNNKNEVHGA
CDVSKRPVDESCSLDPCFYWTITGEMSEVPSEWL"
BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score = 10 Optimized Score = 219 Significance = -0.41
Residue Identity = 18% Matches = 314 Mismatches = 1055
Gaps = 371 Conservative Substitutions = 0
Translation Frame = 1
QLAHGRHQPDLRNAGHIPVSKWGLFYXTTGYGXTRRXRGTKQTPHHLXAQRPPEALNRKACMHLPTQK
10 20 30 40 50 60 70
XAOXROEBNOSKMKRKKXPGMXKRSSTIKORLSNRGIFCLMXKXDOGHKKEPDOKDTFFILSTYCRSLGGGR
80 90 100 110 120 130 140
X MCFVSWATLITLLVLR-----DLAEMSPDAAAVYKRDRLHPROYVLETFSEY
150 160 170 180 190 200 210
QONGFIPIWKRPSTLYNFNFVNQNSIYLXKRWKFKMXYCYCELNDSSXTGWAFHIFPCSDIKTKILVYAFPE
150 160 170 180 190 200 210
50 60 70 80 90 100
EIVSPIRVNALGPEPTN-----VHFRTRRSINSATDPWPAF-----ASSSSSTSQAIYRLSA
110 120 130 140 150 160 170
E--OSRMWPSXYCCSLNKTYGLOSROKXLYLRPGXTGNHLXSLKLELYXXRWLEYSEFYDRPACGVXHA
220 230 240 250 260 270 280
110 120 130 140 150 160 170
EGQGFLENLTANAGFTAPLFTVTLGTPGVNQTFRYSEEDAELEKCFYKGYVNTNSEHTAVISLCSGMLGTF
290 300 310 320 330 340 350
SXHXOQOMXRRRSXESPACHGSNTDELHOPLDVVKYX--KIYHVFPHWLMRVFAKRT-XIQTLPRACPTA
290 300 310 320 330 340 350
180 190 200 210 220 230 240
RSHDGYFTIEPLD-SHDEDEDEEONKPHIYRSAPO-----REPSTGHACDTSHKRHSKKKTRARA
111 111 111 111 111 111 111
R-----HPLQREXTMXIDWTFRPSGPYI--DAYQTALVQXQWQTOR--LPDSAH-----TLG
360 370 380 390 400
250 260 270 280 290
RKMGKINLNLGVAALNS-----GLATEAFSAAYG---KTQNTMEKTKHTRTKFLSPRFVEYL
111 111 111 111 111 111 111
KMDGVNWKALDOWIILCSORNGCPDRDWITGLKLESLMNLDMWRGHNHNSRVQQTFTK-----KWKKIL
410 420 430 440 450 460
300 310 320 330 340 350 360
VVADN-RMVSYHGENLOHYILTMISIVASITYKDPISIGNINIVINLVINHEDDGPISINAOITLKNPCQ
111 111 111 111 111 111 111
CRFXNELYLQGHAMSOAARLRPRATYCS-----LXNEAPYHOSASQCALGYIQMN-----SDGCP
470 480 490 500 510 520
370 380 390 400 410 420 430
WQHSKNSPGIHHNDRVALL--TRQ-DICRAHDKCDTGLGLAEIGTICDPYRSC-----SISDSGLSTAFITIA
111 111 111 111 111 111 111

[illegible]

```

source      1. 5139
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="fh11767"
            /tissue_type="brain"
            /clone_lib="pbluescriptII SK plus"
            /gene="KIAA1312"
            /gene="KIAA1312"
            /note="KIAA1312"
            /note="start codon is not identified."
            /codon_start=2
            /product="KIAA1312 protein"
            /protein_id="BAA92550.1"
            /db_xref="GI:7242979"
            /translation="NSENRYAVISLCSGMLGTFRSHDGYIEPIQSDEDEDEEONK
            PHILYRSPAPREPSTGHAADTSEHNKRSKDKKTPARKWGRINLADVALANG
            LATEFSAVGNKTDNTRKTRHRTKRELSPREVEVLVADNRMSVHG--ENLQHYTL
            TMSIVASIKDPSIGNLINIVNLIVIHNEODGPSISENAOTLTKNFCOMSHKNS
            PGJHHDPAVLITRODICRAHDKDITGLAELGTICDPYRSCSISEDSGSTATTIAH
            ELGHVFNMPHDNNKCKEEGKSPQHYMAPLINTNTYTPMMMSKSKRYITIEDTGY
            ECLLNESRPPIPVOLPGILYVNNKOCLEIFGSOVOCPTIMQCRKLCNNVNGVH
            KCGRTQHTPMADGTECEFGHCKTGFVPEKMDVPVVDGSSWSPGTCSTCGGAI
            KTAIRECNRPRKNGKGCYGRMKFKSCNTPCLAKORDFDEQCAHFGKHENING
            LLPVRYWPKYSGLIMDKRCLFCRVAGNTAYVQLRDRVIDTFCGQDNTDICVQGL
            RQAGCDHVLNKKARDKCGVCGDSSCKTAVAGTENTVHYGYNTVVRILPAGATIDVR
            OSFSGEDDNLVLSKSGEFLNNGNNTYTMARRTIRIGNAVVEGSEIATVERIN
            STDRIEOLLIQVLSGKLYNPVRYSPNIPIDEKPOFTNNSGPMQACSPQCGGR
            KRKLVTRESQULTVSDRCRLPQPGHITTECGDIDLWNVHVSRECSAOCGLGR
            TLDYCAYSRLDKTERKVDGFCSSHPKPNRCKSGECNTGGMVSAETECSSCD
            GGTORRAICVNTENDVLDSCSKTHQKVTIQRCSERPQMKSGDSEELVTCGKH
            KHRQWCOFGEDRLNDRCMDPEKPTSMOTQOQECASWQAGPMGOCSSVCGOYLR
            AVKCIIGFVYVNDNDCNAATRPITDODCEICHPAPAPARTRESTYSGAPRQME
            GSNPSCATCGKTRMYRYSQDENGVSADBSACATTPRYAKRECVTTCGQKALD
            WSSSVTCGGRATROYACVNSDVIIDRSQDODYIPETDQDSMSPCQRPDQSL
            AOHPPQNDYPRKASPSRTHVLGNGMRGPMWACSTAGSGQRVYVCOENGYT
            ANDVERIKPEDQACESGPCQWAVYNGWECTLCGGITRLVYVQNRNGERFPL
            SCETIDRPREDQCNTHACPHDAAMSTGPMWCSVSGRGHROKNVCMKDKDSHLES
            DYCHLAKPHGRKCGRGRCPKMKAGAMSOECTKCGSGYQOARHNGCOIGTHKLAREE
            CNPTPRSEFSCGGRCPRLYTMAEBOECTKCGSGSRIRKAYVCVDDMKNEYHAKR
            CDVSKRPDRSCSLQPCVEYVITGEMSEVPSWEL"
BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score = 8 Optimized Score = 228 Significance = -0.41
Residue Identity = 18% Matches = 337 Mismatches = 1069
Gaps = 369 Conservative Substitutions = 0
Installation Frame= 5
      10      20      30      40      50      60      70
MOFVSWATLLTLLVLDLAEMLGSPDAAAVRKRDLHPROYKLLTETLSEYEIVSPIVNALGEPFTNVHFKRT
      X      X      X      X      X      X      X
      80      90     100     110     120     130
RKSINSATPWPAPASS---SSSTSSSAHTR-LSA-----PQOQFLPNTANAGTAPLFTVTLG
      |      |      |      |      |      |      |
KMYLRYFGYYRKLFSSEAVCYGTGTHLVHERGLFAPYVNLKPKISAPITFKLGNLTYSRFLASLEKSGL--
      30      40      50      60      70      80      90
      100     110     120     130
TPGVNOKTKYSEEBELKHCFTKYVNTNSEHTAVISLCSGMLGTFRSHDGYIEPIQSDEDEDEEONK
      |      |      |      |      |      |      |
---VLDDHISAMPQLEKX-----RCSLHTACASLSQPLSLVLVSLALQYF---LILNCLQ-----A
      110     120     130
      140     150     160     170     180     190     200
PHIIRYRSP-----ORE--PSTGRACCTSEKHN-HSKDKKKTTRAKWG-----RNLG---
      |      |      |      |      |      |      |
PHSLILVCSPCNNFFMPILGHGSEONVDLKGRLDHLHLHSFAFYKEFLXGKAEDLSENTOFVIANRAGSLK

```

```

150      160      170      180      190      200      210
260      270      280      290      300      310      320
DVALNSGL--ATAFSAVGNKTDNTRKTRHRTKRELSPREVEVLVADNRMSVHG--ENLQHYTL
      |      |      |      |      |      |      |
SPTFLSHSLSHLSCHSOMAMELTWT--VPRTPVPLTILX---RHTRRVA---NYSFHPAPACSHSVPH
      220      230      240      250      260      270      280
330      340      350      360      370      380      390
MSIYASTIKDPSIGNLINIVNLIVIHNEODGPSISFNQOTLKNFCOMSHKNSPGJHHDPAVLITROD
      |      |      |      |      |      |      |
APRCCHPRTPPCGTWSLRRSWCI-----LAIPLPKCRGDYVGLSRVPTVWCSTYSLMLCVFERSD
      290      300      310      320      330      340
400      410      420      430      440      450      460
ICRAHDKDITGLAELGTICDPYRSCSISEDSGLSTAFTAHLELG---HVFMPHDNNKCKEEGKSPQH
      |      |      |      |      |      |      |
--SPHASAVLRGHHTQST--DSKRQSLTWIGLITGTCANVAHXGAVSNHFLNGFH-TLPCSHFVYALDH
      350      360      370      380      390      400      410
470      480      490      500      510      520      530
VMAPLNTYTPMMMSKSKRYITIEFDITGYECLLNPE-----SRPYRLPVOLPGILYVNNKOCLEIF
      |      |      |      |      |      |      |
KRONTSSKQGYSMORVYDKHVCYAHDRREVYQEPSSNLETVHRWTADRP---PV-TFLCL-HHTAXCTL--
      420      430      440      450      460      470
530      540      550      560      570      580
GGPSOVCPYMMOCR--RLMCMNV-NGVHKCGRTQH-----TPWADG-----TEDEGKCKYGFVPEKMD
      |      |      |      |      |      |      |
--PCGHPTEDEKGRGHIHLFLFAOVLFSPHSRLRCIHFPDIOQNAAGIRQHRVNNMLPRGQFSTGF-----
      480      490      500      510      520      530      540
590      600      610      620      630      640      650
VPVTDGSGNSPFGTCSRTCGGAIKTAIRECNREPKNGKCYVGRMKFKSCNTEPCLAKORDRPREQCA
      |      |      |      |      |      |      |
HRAHCGGWGRS--GDSPPHGRG-----AELSHCLSEFGMVTWNSPCQFLGN--PGHTHSDRS-----
      550      560      570      580      590      600
660      670      680      690      700      710      720
HFDKRNENINCL-LPNRYWPKYSGLIMDKRCLFCRVAGNTAYVQLRDRVIDG--TPCGQDNTNDICVOGICR
      |      |      |      |      |      |      |
GHCSXHTSLAGLPIRLAHRSSKSS-----SPRPSIAHRYSONIILPLPL--VXAGXHRSRQPSHLSGSXR
      610      620      630      640      650      660
730      740      750      760      770      780      790
QAGCDHVLNKKARDKCGVCGDSSCKTVAGTFTNVHYGYNTVVRILPAGATIDVRQHSFSGETDDNYLA
      |      |      |      |      |      |      |
ISSGYLSHKWLSMGSKTQIATGFLVHCMSFSPGOLGEDDMVYHSP---GYOLVXLHYSCHNPLRY--
      670      680      690      700      710      720      730
800      810      820      830      840      850      860
LSSKGEFLNGNVVTMAKREIRIGNA-VVEYSGSETAVERINSTDRLEQEL--LLQ-VLSYGKLY-----
      |      |      |      |      |      |      |
KSQXQISLLLAGIIVNHKXCTVPRDPRARMHIRPADKSAXMLMSOGNFTYHLLIDHOTRPGGAYAFHNR
      740      750      760      770      780      790      800
870      880      890      900      910      920      930
NPVRYSFNPIPIEDK--PQGFYWNSSHQWQACSKP-----CGEERKRLKLYTRESQULTVSDRCRLP--Q
      |      |      |      |      |      |      |
XPSITLQISTVKGCTNCTSEMXLSDVLCICCPVNHFGYKXKPFSSSHHSLSLNIQSRONSASHPCYI
      810      820      830      840      850      860      870
930      940      950      960      970      980      990
PGHITTEPC--GTDCDLRWNVHASRECSAOCGLGYRTLDITCAKYSRLDGTETKRVDDGFCSSHRKPSNRKRC-
      |      |      |      |      |      |      |
PLSITFHGCLVWDGCGCKNHQPSOSHPACYIMHSR--CPMGNSHTGHYIHSCWOAHATSGHLYRVQXCV
      880      890      900      910      920      930      940
1000     1010     1020     1030     1040     1050
---SGECNTGWRYSANTECSSKCDGSGTORRAICVNTFRNDVLDSCKT-----HOKRYVTIQR-CSSEFP-
      |      |      |      |      |      |      |
QAGAADRIFOKOXADQIPWCKQOVFAVSPLGRVYC--MLAMADCSSTAEVYLDQLEYKKNVSRGCTTFPP
      950      960      970      980      990      1000     1010

```

1060 1070 1080 1090 1100 1110 1120
 -CPWKGSDMSECVTGTGCKHKNRWVWCGEDRLNRMCDPEKTPSMOQOPECAWAGWGGCS---
 TKPRAKVLACDIL-----SKFELRQ-----SRTHCLLPQ-HSCCEPFW---PLXOOSFHL
 1020 1030 1040 1050 1060
 -----VTCGGGQYR--AVKCIIGYMSVVDNDNC--NAATRPIDTQDCELPSCHPPAAPETRRS
 ARTHLYCLIKSSCHRLPRLNCAAAQWYXMHOLEFGROYCNHVLXXMFLPFCMKNYRHHKPHILSLGL
 1070 1080 1090 1100 1110 1120 1130
 1180 1190 1200 1210 1220 1230 1240
 TYSAPRTQWRFSGWTPCSATCGKTRRMYVSC---RDENGSVADE--SACATLPRPAKECCSVTPCGOWKA
 SLKHDRIQ-----LAGKPRGRHYHLCSPGKKEFHSLICEADSRILCSLP-----LCRTTCG-
 1140 1150 1160 1170 1180
 1250 1260 1270 1280 1290 1300 1310
 LDMSSCSVTGCGG-----RATROYCVNYSVDHYIDRECDQDYPETDODCSMSPCEQRPDPDGLAQH---
 --PSS-EFHCIYXGSAHMEADR--XCNNAS---RQSE-HTVHLGSLASAXDMAPC-CRTYISFYVLSIFH
 1190 1200 1210 1220 1230 1240
 -----PFQNEHYRPRRSASPSRTH-VLGGNQMWTGPGACSSITCGAGSORRVVVC
 HFLVLCSTLEMLEXCPHLMFWRSRFDSDNPPR-IHLSRGHPFLMEKHRT-----CSAFQARTPSRPRVC
 1250 1260 1270 1280 1290 1300 1310
 1360 1370 1380 1390 1400 1410 1420
 ODENGYANDCVEREIKRDEORACSGPCPOWAYGNWGECKTLGGGIRTRLVVCORSNGERFPLSCILLDK
 -AESG---SLCV-----LHKRYCTRAVC-TASYMGTEPNLYOKSIHIVYSRCKCLAVGAQKGRVWIOVRXA
 1320 1330 1340 1350 1360 1370
 1430 1440 1450 1460 1470 1480 1490
 PPDEOC-NTHACHDAAMSTGPMWSSVSCSGRGHKORNYCMKDKSHLESYD---CKHLA-KPHGRKCR
 NTRNOCILKTOXYTFDYTLTT-----SRGMCRSSVLEFXHAGD-SXLLLLATYCCCHHEACXTHG---PA
 1380 1390 1400 1410 1420 1430
 1500 1510 1520 1530 1540 1550
 GGRCPRKAGAMWSSCSVSCGRVQORHVGCOIGTHKTAARETECNPTTRPE-----SERDC-----Q
 HGRXKLYSITHYLNKXMSFYRDHKMFVOPGLRYH-ICRELGRYPVLLREDOYHGDGPHLLDCSSNAATGKSFLL
 1440 1450 1460 1470 1480 1490 1500
 1560 1570 1580 1590 1600 1610
 GPRCPLYTMR-----EEMQ-----ECTKTGEGSRYR-KVVCVDNKNVHGANCDVSKRPVDRSCSLQPC
 MLSEHKKIMRANRPNHYESOLSSQXOYLNFQYLGILYXRJLDLTLLKYSEVGFHGMKRFCCLPRLLO--
 1510 1520 1530 1540 1550 1560 1570
 1620 X
 EYWTIGEMSEVPSWEL
 TVDRIKYNLWSSGSSFLPCCPSYYHKKMPLLSRCLMLLRHOPGXSLPFIILLMFSSCLYCAFYCVLRCHM
 1580 1590 X 1600 1610 1620 1630 1640
 HAFLLRLSGGRCAKXKGCYCFVPLHLVHPXYVVOXNNPHHETGMCAPRALISGXKXPPCAR
 1650 1660 1670 1680 1690 1700 1710

4. US-09-972-467-2 (1-1629)

AB037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733

VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_id:pbuescriptII sk plus
 clone:fh11767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (sites)
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M., and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 DNA Res. 7 (1), 65-73 (2000)
 JOURNAL MEDLINE
 20181126
 2 (bases 1 to 5139)
 Chazara,O., Nagase,T., and Kikuno,R.
 Direct Submission
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@info.kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 FEATURES
 source
 1. 5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh11767"
 /issue_type="brain"
 /clone_lib="pbuescriptII sk plus"
 1. 4417
 /gene="KIAA1312"
 <1. 4417
 /gene="KIAA1312"
 /note="Start codon is not identified."
 /codon_start=2
 /product="KIAA1312 protein"
 /protein_id="BA092550.1"
 /db_xref="GI:7242978"
 /translation="NSEHTAVLSLCSGMLGTFPSHGDFYIEPLQSDMDQDEDEEONK
 PHLYRRSAQRPESTGRACDTSEKRNRSKDKTKTRAKWERIYNLADVAALNSG
 LATEARSAGNKDNTNREKTRHRTKRLFSYPRFVLLVADRWVSYHGENQHYIL
 TMSIYASISYKPSISIGNLINVIVNLIVIHNEGDPSISNAQTTLKNCOMOHKSN
 PGSIHNDTAVLLTRDIDICRAHDKCDTLGLAELGTICDPYRSCSISDSGLSTFTLH
 ELGHVFNMDPNKCKEKGVSPOHVAPTLFTNPPMMSKSKRYTFEFLDTYG
 ECLINPESPSPYPLPOLDGIILYNMKOCELITGSCQAPYMKOCCRLMNVNVC
 KGRTOHTFPAWDTCEBPGRHCKTGVCPREMDVPLTDSGWSGSPFGTCSRTCGGI
 KIAIRECNPREPNNGKYCVGRMRKSCVTEPCLOKROFDEQCAHFDGKHFNING
 LLEPNRMVPRYSGLIMKDRCKLEGRVAGNTAYOLDRYIDTGPCQDNDILCOGLC
 ROAGCDHVLNSKARDKCGVGGDSSCKTAVAGTFNTHVGYTVVIRIPAGATNIDVR
 OHSEGETDDNDYALSSKGEFLNGNFVIMAKREIRIGNAVVERISSEAVERN
 STDIRIOELLQVLSGKILNPDVRSFNPIEDRQOQPYMNSHGWMQSKRCQGER
 KRLVCTRESDDQVTSDDRCDRLPQGHITTEPGCTCDLKNHVASRSECAQGLTR
 TLDLYCAKYSRLDGKTEKYDDGSSHPKPSNNEKSGDEGNTGMYSAMTECSKSD
 GGTORRALICVNTNRNDVLDSCTHOEKVTIQCSEFPQWKSMDSELYCGGGLG
 KHRVWMCQFGEEDLNDRMCDPEKTPSMOQOPECAWAGWGGCSVTCGGQYOR
 AYKCIIGYMSVVDNDCAATRPIDTQDCELPSCHPPAAPETRSTASAPTRRF
 GSWTPCSATCGKTRRMYVSCRDENGSVADESCATLPRPAKEECVTPCGOWKALD
 WSSCSVTGCGGATRQWCVNYSVDHYIDRECDQDYPETDODCSMSPCEQRPDPDGL
 AQHPFQNEHYRPRRSASPSRTHVLGGNQMWTGPGACSSITCGAGSORRVVVC
 ANDCVEREIKRDEORACSGPCPOWAYGNWGECKTLGGGIRTRLVVCORSNGERFPL
 SCEILLKPRDQECNTACGPHDAAMSTGPMWSSVSCSGRGHKORNYCMKDKSHLES
 DYCKHLAKPHGRKCGRCPRKKAAMWSSCSVSCGRVQORHVGCOIGTHKTAARETE
 CNPTRESERDQGPRCPLYTMRAREMEQCTKTGEGSRYRKYVVCVDNKNNEVHGAR
 CDVSKRPVDRSCSLQPCFYWTIGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803

Initial Score = 8 Optimized Score = 80 Significance = -0.41
 Residue Identity = 104 Matches = 112 Mismatches = 333
 Gaps 3
 Translation Frame=

```

TVQVESIMILLESYQDHSLEITNVIPXAMLNKPEVPIEAVLAVKIVDXVQLRSPMSAMCITCLMMT
220 230 240 250 260 270 280
TTNNKKELRYVPSMSLQHXSTPTPCGQGVVENISLSFXTLVMAVSCLTNLNPDTLCLSCQASFTTXI
290 300 310 320 330 340 350 360
NNNKKFLDVOYLRCHAYCSADGSGAITSMEYTKAAGLSTHPGMRSGASLESTASMDVEFPKKMSPXMDP
370 380 390 400 410 420 430
GEVGVPLEPAPEHEVGAQKQPFESATDONOKMVENIVXDEYNLSPATRSVHSRSSETSMSNVLTITGSL
440 450 460 470 480 490 500
TSTYCFPMGAGSLNTVEFYXKRTGASCAENOGTOPTISFETEXAMELLVARTOMISVSAFAGKLDAIMEXT
510 520 530 540 550 560 570
OKPGEINGVFVAIILHAQWQEHILQYIMVTLIMSEFQVLPLIMCGSTVSOGKQMTTXXLYQAVKNVSC
580 590 600 610 620 630 640
XMETLLSOWPKGKRALGMLKXSTYGPRLPKKELTQOIALSKNCFRCRRESCTTPMVAIILSIFOLKINSS
650 660 670 680 690 700 710 720
FTGTVMGHGKHAIVNPAKNGNENLFAFGNLSLFLIKDAIGCPBLDTLNPVQVTVTXGMLPAGVNVPS
730 740 750 760 770 780 790
VAWYTAHMTSTVPIAGMGRRLRLMVAVFAAIPNOATVKNAGONTVAGAILPGLINQKAVTVGPREGEL
800 810 820 830 840 850 860
FVSIPEMMYMTANAHIKRRLPERGAVSLVHSGNLETGOSAWPVEKISTARSGSLVXIKIDMIETCYTLR
870 880 890 900 910 920 930
PSOHLCLYSSRNHPRARYPGSAVSLVDRDTSXEQXNASLGLICOMKMTNVVMOQLOLIRTVNHHVI
940 950 960 970 980 990 1000
LPOLPRRGEAHYHQEPSSDGLIGPRAOPLVGRVPGXDTSAEMRMALVTLRVPLPCLDQOKRKNVLXHP
1010 1020 1030 1040 1050 1060 1070 1080
VGNGRPWTGALALXPGKVGQPGKXCSTVTVTXSIGVSVTRILISOKLTFTVCHHALKPGQVAXLSTPSC
1090 1100 1110 1120 1130 1140 1150
MRTIVGAPAPAPMCSVETSGELAPGEHYVPVPLADPSGVLLYEMKMDTPOTVYREXNLMSKEPVNPA
1160 1170 1180 1190 1200 1210 1220
10 20 30 40 50 60
LTLTLVLDLAEMGSPDAAA-----AVRKDRLHPRQ--VKLETLSEYEIVS-----PIRVNALGE
11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```

```

LRYAESXVEGVQEOCGLEATPRICHCELCVEFXOYLSPF-KLEFLVCKAGENXENXVSKLOAYILXPTLMS-
1480 1490 1500 1510 1520 1530 1540
VASTYDPSIGNLINIVYINLVYIHNEODGPSISFNAQTTLNKFCQ-WQHSKNSGGLHHDTAVILLTQ-DI
--YMHKEEITGAAN---SXDAVRCILKIVYKRIILXCOGYNNERQLOXTSTCYKQ-----AESLSLOLWP
1550 1560 1570 1580 1590 1600
400 410 420 430 440 450 460
CRAHDKDPLGLAIEIGTIDDPYRSCSISDSCLSTAF-TIAHELCHV-NMFDHNNCKKEGKYK--SPQHY
1610 1620 1630 1640 1650 1660 1670 1680
CRNVYRYQTFFKRXKSGGLCEISXLENGXNFMFXNIVYKGTQTPPMQWCCALTNFRGRKFSIIE
1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
MAPTLNFTYNNPMMSKSCSKYITTEFLDTYGGCLINPEPSRYPPLVQPLGLILVYNKQCELIIFPGSQVCP
VSOVPE-----HISCAQCLSRPFL-----MVINQXSS
1960 1970 1980 1990 2000
540 550 560 570 580 590 600
YMQQCRMLCANNVNCYHKKCRTPQHTPMADGTCEPFGKCKYGFYCPKEMDVPYVDGSGMSMSPGTGCRTCG
610 620 630 640 650 660 670
GGIKTAIRECNRPPEPANGKCYCVGRMRKFKSCNTEPCLKOKRDFRBOCAHFQKHFENINGLILNVBRWPKY
680 690 700 710 720 730 740
SGILMKDRCKLCRAVAGNAVYQLDRVIDYDTPGCGDINDICVGLCQACQADIVLNKAKARDCKGCVGSDN
750 760 770 780 790 800 810 820
SSCKTVAGTFNTVHYGYNTVVAIPAGATINIDVRQHSFSGEDDDNYALSSKGEFLINGNFTVYMAKREIR
830 840 850 860 870 880 890
IGNAVEYEGSSTAVERINSTRIDIEQLLOVLSYGLKLYNDVRSFNIPLEDKRPQCFYWNMSHPWQACSMP
900 910 920 930 940 950 960
CGGERRKRLVCTRESQDLTVSDQRCDRLPQREHITTEPCGTDODLHMVHARSSECSAQCGLGRTLDIYCAKY
970 980 990 1000 1010 1020 1030
SLUDKTEKVDGFCSSHPKPSNRKCKSGECNTGWRYSAMTEKSKSGDGGTQRRRAICVATVTRDVLDDSKC
1040 1050 1060 1070 1080 1090 1100
THQEVVTLORCEFPCCPOWKSQDWSSECLVTCGKHKRWQVOCQGEDELRLNMCDPETKPTSMOTCOQPECA
1110 1120 1130 1140 1150 1160 1170 1180
SMQAGPWCQCSYTCQGYQLRAVKCIIGTYSVVDNDNCNATRPDTQDCELPSCHPAPAPETRRSTYSA
1190 1200 1210 1220 1230 1240 1250
PRTQMRFGSMTFCSATCGKGRIMRKYSCRDENGSAVADBSACATILPRPYAKKECSYTPCGOMKALDMSCSYV
1260 1270 1280 1290 1300 1310 1320
CGQGRATQVMCVNVDHVIDRSECDQDIPEITDDCSMSPCQRTDPSGLAOPHFQFQNEEDYRPRSASPSRTH
1330 1340 1350 1360 1370 1380 1390
VLGNGWMTGPMGACASSYTCAGSGQRRVYVCCODENGYTANDCVENIKRPEQACAGSGCPQMAIYNMGCEIKL
1400 1410 1420 1430 1440 1450 1460
CGGRTIRLIVVCOBSNGERFPLDSCCELLDKRPDRQCNTHACPHDAMSTGPMSSCSVSGRHKORNVYCM
1470 1480 1490 1500
AKDGSHTLSDYCKHLAKPHGRKCKGRCRPMK

```

5. US-09-972-467-2 (1-1629)

ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978

KEYWORDS Homo sapiens brain cDNA to mRNA, clone_1lb:pb1uescriptII SK plus

SOURCE clone:fh11767.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.

TITLE Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro

JOURNAL DNA Res. 7 (1), 65-73 (2000)

MEDLINE 20181126

REFERENCE 2 (bases 1 to 5139)

AUTHORS Ohara,O., Nagase,T. and Kikuno,R.

TITLE Direct Submission

JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:htp://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914)

FEATURES

source location/Qualifiers

1..5139

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="fh11767"

/tissue_type="brain"

/clone_1lb="pb1uescriptII SK plus"

1..4417

/gene="K1A1312"

<1..4417

/gene="K1A1312"

/note="Start codon is not identified."

/codon_start=2

/product="K1A1312 protein"

/protein_id="BAA92550.1"

/db_xref="GI:7242979"

/translation="NSEHTAVISLCSGMIGTFRSHDGYFIEPLQSMDEDEEENK PHILYRSAPREPSTGRACDTSEKRNRSKDKTKTRAKWGERINTLAGDVAALNSG LATEAFAYGNKTDNTRERKTRKFLSYPFVVLVADNRMYSGENLOHNTL TMSIVASYIKDPSIGNLINIVNLIVIHNEGDGSIENAOITLKNFCOMOHKNS PGCIHHTFAVLIRODICRAHDKCDTLGLAETGIDDPYRSCSISPDSTAFATTAH ELGHVEMPHDDNNKCKEEGVSKQHYMAFTLFTNPMWMSCKSKRYTTEPDTGAG ECLNPEPSRPPLPVOLPGILYVNMKQCLIFGPGSYVCPIYMOQRILMCNNVNVH KGCRTQHTPMADTECEGPKHCKYGCVCPEKEMVPTLDSWGSWSPFGCSKRGGI KTAIRECNREPNKNGKCYGRMKPKSCNTEPECLKOKDFRDEOCAPHDKHFNING LLPNRMVPSYSGILMKDRCKLFCVAGNAYAYOLDRVIDGPGODINDICVOGLC ROACGDHVLNSKARRDKCGVCGDNSCKYVAGTENTMYHYGVYRIPRAGNTIDVR OHSGSGETDDNTLALSSSKGEFLNGNFTVYAKREIRIGNAVYVSSSETRAVENIN STRIDIEQLLOVLSVGLKLYNPVRYSFNPIEDKPOQFYWNHSGWQCSKPCQGER KRLVCTRESDDQLVSDORCDRLPOGHITPEPCDDCLRWNASRSESAQCGLGR TLDIYCAKYSRLDKTEKVDGFCSSHPKSNREKSGEENGTGMYSAMTECSKGD GGTORRAICVNMTRNDYLDKSKTQHEKVTIORSEFPCKPMKSGMSESLVTCGCH KHRQWVCGEEDRLNDMCDPEKTPMTSMOCCOPCASMGAGPMSGCVTCGGVYLR AVKCIITGYSYVDNDNCNATRPPTPODELSRCPFPAPRPTRSYTSARTQWRFS GSWTPCSATCGKGTMRKYVSCRDENGSADESAALPRVAAEBCSVTPCGWKLAD WSSCSYCGGGRATQWCVNYSNVSDHV IDRSECDQDIIPETDCCSAPCPORPDSGL AOHFQNEIDYRPRASRTHVLGDMOWRTGPMGACSTCAGSGSRVRYVCDENGT ANDCOVERIKPDEORACEGSPQWAGNMGECKILGGGIRLVLVCCSNGEPRFDL SCETLKKPRPDECNHACPHDAWSTGSCVSSGCHQORANYCMAKGSNILES DYCKHLAKPHGHKCKRGRCRPPKAKAMSCSVCSCGRVQOQRHVGCOITHTIARTE CNPYTRRESRDCQGPRLTYWRAEWOECTTTCGEGSRKYRVVVCVDNKNKEVHAR CDVSKRVRDESSCLQCEYVYITGEMSEVPSEWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..

Initial score = 8 Optimized score = 154 Significance = -0.41

Residue identity = 19% Matches = 220 Mismatches = 668

Gaps = 249 Conservative Substitutions = 0

Translation Frame= 4

```

10      20      30      40      50      60      70
MOFVSATLTLTLVLRDLAEMSGSPDAAAAYRKDRLPROYKLLLETISEIYSPITRVNLAGFEPTNYHFRT
80      90     100     110     120     130     140
RRSINATPMPWPAFASSSSSTSSQAHYRLSAFGQFLFNLTANMGAFTLPTVLLSTPGVNTKPYSEEB
150     160     170     180     190     200     210
AEIKHCFYGYVNTNSEHTAVISLCSGMIGTFRSHDGYFIEPLQSMDEDEEENKPHILYRSAPQREP
220     230     240     250     260     270     280
STGRHACDTSEKRNRSKDKKTRAKWGERINTLAGDVAALNSGLATEAFSAVGNKTDNTRERKTRKRF
290     300     310     320     330     340     350     360
LSTPFEVEVLVADNRMYVSGENLOHTLITLMSIVASYIKRPSISGNLNIYVLYIHNHODPSISFNA
370     380     390     400     410     420     430
OTTLNFCOMOHKNSKPSGCIHHTDAVLIRODICRAHDKCDTLGLAETGIDDPYRSCSISEDGLSAFTLI
440     450     460     470     480     490     500
AHELGHVEMPHDDNNKCKEEGVSKQHYMAFTLFTNPMWMSCKSKRYTTEPDTGAG ECLNPEPSRPY
510     520     530     540     550     560     570
PLTVOLPGILYVNMKQCELIFGPGSYVCPIYMOQRILMCNNVNGHKGCRTOHTPMADGTECEPEKHKYGF
580     590     600     610     620     630     640
CYREKMDVPYTDGSGWSMSPFGCSKNTCGGGIKTAIRECNREPNKNGKCYGRMKPKSCNTEPECLKOKRD
650     660     670     680     690     700     710
FRDEQCAHFDGKAFENINGLLPNRYWPKYSGLIMDKRCLFPRVAGN-TAYVOLDRVIDGTPCCODINDIC
680     690     700     710     720     730     740     750
F-----FLXSCLL-----NRHNTSGSKGFVCPLOCFKTKNESPHQVYKSHI-LVQIFLSXKI
760     770     780     790     800     810     820     830
VOGLCR-----OAG-CDHYLNSK-----ARRDKCGVCGDNSCKTVAGT-----FNTHYGY
780     790     800     810     820     830     840     850
RSGNAPHFQMATAGYKELLILSHSKXEVTVASIPCCIPGTTIFSYTLSSGTSQHPMSLOPLXFFLYAN
860     870     880     890     900     910     920     930
NRYVRIIPAGATNIDVYOHSEFGTDDNTLALSSSKGEFLNGNFTVYAKREIRIGNAVYVSSSETRAVER
940     950     960     970     980     990     1000     1010
RTSRXITTECRPESTXPSLTSILPCLIOGILAMBS-----XGLVRKHTVHSGK-----SGGXQVSH
1020     1030     1040     1050     1060     1070     1080     1090
840     850     860     870     880     890     900     910
INSTRIDEQLLOVLSVGLKLYNPVRYSFNPIEDKPOQFYWNHSGWQCSKPCQGER
1020     1030     1040     1050     1060     1070     1080     1090
IPALIPPLPLPLP-----SGHXADD-GYS-----SODGTSDSHPVIOYVSQCKLQLSRSTGRLLTSO
1100     1110     1120     1130     1140     1150     1160     1170
220     230     240     250     260     270     280     290
910     920     930     940     950     960     970     980
R---ESDQL--TVSDOR-CDRLPOGHITPEPCDDCLRWH---VASRSECAQCGLGRITLDI-----
1020     1030     1040     1050     1060     1070     1080     1090
RACSTFLLSTTTTTRLYEPPQVVLVHSCHSALOVXGHHGPMQSRSDSLVYGLASVSALILCVPIXP
1100     1110     1120     1130     1140     1150     1160     1170
280     290     300     310     320     330     340     350
970     980     990     1000     1010     1020     1030     1040
YCAKYRLDKTE--KVDDGFCSSHPKSPNREKSGEENGTGMYSAMTECSKSDGGTQRRRAICVNTR--
1100     1110     1120     1130     1140     1150     1160     1170
TCACCTPRPDQTEHXLQADAFNIGHLP-----RHFLCPGGLAR-CLQXLSLKWLPSPFAMQXTFCLCPRPQ
1180     1190     1200     1210     1220     1230     1240     1250
350     360     370     380     390     400     410     420
1030     1040     1050     1060     1070     1080     1090     1100
-----NDVLDSKCTQHEKVTIQRCS--EPPCPQWKSGDMSSECLVTGCKHNRQYWCQGEDRLNDRMC
1120     1130     1140     1150     1160     1170     1180     1190
ETHEHLOGPVLHAAASC-GQACVLYHCSRSGLSRISQLKSGNRS-----PLDRXQTTSL
1200     1210     1220     1230     1240     1250     1260     1270
420     430     440     450     460     470     480     490
1100     1110     1120     1130     1140     1150     1160     1170

```



```

DEPRTSMOTQOPEECASMOAGPWG--CSYTCGGYQLRAVKCTIGTYNSVVDNDCAATPTD--TGD
VLMPPHSVLHSPOLPXAHXGQPDSDALCSSGFLSTQSA-----VYFSSXHTTTRMDPAV
470 480 490 500 510 520
CE-----LPSCHPPRAAPTRSTYSAPRTOMRFGS-WTPCSATCGK-----TMRVYSCDENGSADESA
LEHAPQGVLEHWP--PSTWRLGLALR--GRXSFMKGCAKPLSGVLXGHEQSWSVSGIXSMHSLR
530 540 550 560 570 580 590
1160 1170 1180 1190 1200 1210 1220
CATLRPAVAKKECSY-TPCGOMKALDMSSCSVTGCG--RAT--ROYMCVNTSDHVTDRSECODYIP
1230 1240 1250 1260 1270 1280
SITTSALATHTTCVVALDPQVTEOELOSKAFHCPQGVTEHSSFPATGLRVAQALSSATEPSSROLTYLI-
600 610 620 630 640 650 660
1290 1300 1310 1320 1330 1340 1350
ENDQDCSPSPORPDPDGLAHPFONEDYRPRASPSRTHVLGNOVRTGPMKACSTCAGSGQRVYVCQ
-----RVPPRQ-----VAHGQY--DPNRHMYGALYVLLVSGAAGGXDGNSSGWSVGLVALQ
670 680 690 700 710 720
1360 1370 1380 1390 1400 1410
-----DENGTYANDCYERIKRPDQ--RACSGCPOMAYGMGECTKRLGGGIRTRIV-----VQDRNG
SLSSSTDLXVMMHFTALSMWPCPOVYLHCQGPACDA--HSGCXQVC--IDVGLSGSHILSPNRSSP
730 740 750 760 770 780 790
1420 1430 1440 1450 1460 1470
ERFDLSC-----ELIDKRPD--REOCNTHACPHDAAMSTGPMSSCSGSGHGHKORHNYCMADGSHL
NXHQTWRCLCFPOVYTHSDSPFHCGGGSLLHLMVTPSCXVHLLSSTS-FRVLTQ-IALLVWPBQL
800 810 820 830 840 850 860
1480 1490 1500 1510 1520 1530 1540
-----ESDCKHLANP-HGHRKCRGRCRCPKAKAGAMSOCSGCGVGOORHNGCIGTHKTARETECNPY
FEHVSQAEHRQRPVLHSEHFSRLGLGMLQKSTSFVSFPPSLYLQAXMSVNRKPHH-WALHSLLAT
870 880 890 900 910 920 930
1550 1560 1570 1580 1590 1600 1610
TRPSENRCQG-PRCPPLYTWRAEMOBCITCTCGEGRS-YRKVVCVDNDKNEVHGARCVDKRPDRSCSLQ
CHLRSQVPOGQVACP-----GWSGRSHLXSETVASDSDLVQTSFRRSPMOGLLHACGIP
940 950 960 970 980
1620
PCEYVWITGEMSEVPSWEL
X
1630
XLFQXNCXGLSTIGILEKXRTSGLYNFTDRTKXSSCSMSRYELLSTAVSDPIYSTTAFPMRISLATV
990 1000 1010 1020 1030 1040 1050
TKRFFKNSPLLDKAXXSSSVSEKLCRTSILVAPAGIRTTVLXPKCTVLNVDPATVYLHEELSPQDPHL
1060 1070 1080 1090 1100 1110 1120 1130
SLRAFEKTXSHRACRQRPWTQISFVSWPQGVPSITLSRSXXXAVPATLQNNLHRSFIRIPLYLGTORTLS
1140 1150 1160 1170 1180 1190 1200
SRPLMLKCFPSKXAHGCSRSRCLRHGSLVDLNLTLRPTQYFPPFPFSGGLLSHSMAYLMPPHVLEOVPK
1210 1220 1230 1240 1250 1260 1270
GLQLPQDPSTVGTGISISLTQNPYLQCPGSHSVPSAQGVXVRPCLPTLLIHOSRLHCITTYGHHXEEGPK
1280 1290 1300 1310 1320 1330 1340
1350 1360 1370 1380 1390 1400 1410
INSICLFTLXRMGSGWTGKXGIDSGLSKHSXPVSKNSVIYFRLHFHDIQGLVXKFSVAGMTGMLTPS
1420 1430 1440 1450 1460 1470 1480 1490
SLHLLLSXGMLTWPSSNAIVKAVLNPISLLEOLLXGSOAWPSSARPPVSHLSXALQISCLVKTATVAXM
1500 1510 1520 1530 1540 1550 1560
IPGGLPECCCHQKFRNVYVALKDMGPPCSLXITTKFTTTLIKFPILGSLXIEATIDIKYKIXCXRPSPW

```

```

YETILSATTKTSSNRGXDKRFFVLLWLESLVSLVLPXAEASVAKPLFNAAATSPARLLSPHFLALVF
1570 1580 1590 1600 1610 1620 1630
LSLCLFLCSGVSHACIPVEGSLMGALRLXMMWGLFCSSSSGSSIDCSGSIKXSPSDRNPVPSIPEORLMT
1640 1650 1660 1670 1680 1690 1700
AVCSE
1710
6. US-09-972-467-2 (1-1629)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b.pbluescriptII SK plus
clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirotsawa,M. and Ohara,O.
Prediction of the coding sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
20181126
2 (bases 1 to 5139)
Chara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@info.kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
source
location/Qualifiers
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/tissue_type="brain"
/clone_lib="pbluescriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="KIAA1312"
/note="Start codon is not identified."
/codon_start=2
/product="KIAA1312 protein"
/protein_id="BA082550.1"
/db_xref="GI:7242979"
/translation="NSETHTAVISLCSGMLCTFPSHDGYFIEPLQSDMEDEEERONK
PHILYRSAPQREPSSTGRHACDISEHNRHSKDKTKRAKMERGLNAGDVALLSG
LATAPFAVGNKQDNTREKTRHRTKREFLSPPFVLYLVANRMYSGEENLQHYLT
TIMSVASITKDSIGMLTIVLVNLVTHNEDGSSISFNATQITLKNRCOMHSHNS
PGGIHDTAVLRLRQDTCRAHDKCDTLGLDELDTCDPYKSCSISDSGLSTAFTHA
ELGAVFEMPHDDNNKCKEEGVKSPQHYMAPTLNFYNNPMMWSKCSKRYITFELDTYG
ECLINDESRPYPLPYDPLGILYNNVKOCELIDPGSOVCPYMWQCRRLMNNVNRVH
KGRTOHTPWADTECEPGRKHCYGEFCVPEMDVPTDGSWGSMPFGNCSRRGGGI
KTAIRBGNRPKEPNNGKCYVGRMKCSCTEPCLOKXDFRDEQAHRDGKHFNING
LTPNVRVRYPSGILMKDRCKLECRVAGNATYQLDRVIDTGPCGDTINDICVQGLC
ROAGCBVLSKARRDKCGVGGDNSSCTVACTENTVHGTYVVRIPAGATNDIVR
QHSFSGETDDNDLALSSSKGEFLNGNEFVYTAKEIRIGNAVVEYSSGEAVERIN
STDIRDEBLLOVLSVGLKLYNPVRYSPNIPIDKQPOFYMNSHGFWQACSKPCQGER
KRLVCTRESDDLTVDQRCDRLPQGHGTEPCGTCDDLRMHYASRSECSAOCQYER
TLDIYCAKYSRLDGKTEKYVDGFCSSHPSPNRNEKSGECLNGGMYRYSATWTCSSCD
GTGRRALICVNRNDVLDSDSKCTHQBKTYIQCSFEPFCPOKMSGHSECLVTCGKGH

```

KHROYWCOGEGEDRLNDKRDDETRKPTSMOTCOQPEPCASQWAGPMSGCCSVTCGGCYOLR
AVKCIIGITMYSVVDNDNCNAATREPTDODCELPSPAPAEETRSTYSAPRTOMRE
GSMTPCSATCGKGTMRMYVSCRDENGSAVDSACATTLTPPAKECSVTGPGOMKALD
MSSCSVTCGGGRATROVMCVNSDVIIDRSBODDIYETDODCSMSPCORTDSGL
AOPHONEDYRBRASPSRTIHLGDNOMRQWAGSSCTCAGSQRHVCOCDENGY
ANDCYERIKRDRORACEGSPCOMAYGNMGGDEOTKLGCGIGRTIPLVOCORSNGERPD
SEETIDPRDQOCYTHHACPRHDAANSTGPMSCSVCSCGRKHOKRVNYCMAGSSHLES
DYCKHLAKPHGRKCRGCRPKMAGAWSQCSVGVQOQHNGCOIGTHKIARETE
CMYTRPESERDCGGRPCRELYTWRAEMOECECTKCTCGSGSRKRYKVCYVDNKNNEVHGAR
CDVSKRPVDRSCSLOPCEYVITGEEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t

ORIGIN

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score = 7 Optimized Score = 205 Significance = -0.41
Residue Identity = 18 Matches = 303 Mismatches = 999
Caps = 325 Conservative Substitutions = 0
Translation Frame= 6

MOFVSMATLLTLVRLDAEMGSPDAAAVARRDLRHPROVKLEETSEYELVSPDIRVALGEPPRTVWHFKRT

RRSTNSATDWPAPAFASSSSSTSQAHYRLSARQOFLFNLTANAGFIAPLTVTLGLTGPVQNOTFYSEE

150 AELKHCFYKGVNTNSEHTAVISLCSGMLGTFRSHDGYFI---EPLOSMDQED-----EEEN 200

150 DYWFITIKGLDKCAQEDIKKNGTXXDSDIENYF X 10 20 30

210 KPHILYRSARQREPSGTRHACDTSEKHNRSKDKKTRARKWGERINLADVAALMSGATEAFSAVYKNT

40 LPLILFYAQAHIMFMKGVCPLTF-----MFXNQKFOPPSSXKESHTSPDEXPL-----LKNQVWX

280 DNTSEKTRHRRTRKRFSLYPRFEVLYL-VADNRMYSYGEMLOHYILTLMSIVAS-LYKDP-SIGMLINIVI

100 RXTFELHSHSESR-DSAPRTOHVLVCHSCHLSILLYPMHYNIFLYLIVFRLHLSYEFAPRVIISICQXDI

350 VNLIVIHNEODGPSIS--FNAOTT-----LKNFCOWHKSNSPGIHNHTAVILLITDODICRAHDKCDTL--

170 KVNHRMXXRRLETTISYSPALHTRNSSLGKLRFCOKTHS--SOMOI-GRVASSLPSTCHTNPSTHPSATLRP

410 -----GLAELTICDPYRSCSISEDGSL-STAFITIAHGLHVFNMN-HDDNNKCKEGVKSPOHVA---PT

240 WLSXWMLQFPBRVYLXP-FSCDPLIAGIQTATFVHPRLAHVTAQPMHVLVYVHTHHLAVPGAFAGLGA

470 LNFYNNPMWMSKCSKRYITTEFLDTGYEGCLLNEBESRPPYLPVQLPGILYNNVNOCEL-----IFPGSQV

310 LPFLCPPEVEGPMALAVAFRLRSGVWVALGLSGFVCSDLTAMPLLYLASATGHRALTSPASPFASS--

540 CPYMMQCRRLMCNNVNGVHKGCRTOHTPMADGTECEPGKCKYGFVCPKEMDVPTVDSGWSMSPFGTSRT

380 --SSALSTPMMLSOVLVITFXMASIFCHAVNISLPMSTSTRDRTRAP---RASTCSVYVMTSMCVT-LLT

610 CGGGIK--TAIRECNRBPCKNGKGYCVRMMKFKSCNTEPCLKQRDRFRDQCAHAFQCKHFNINGLLPNRW

450 IGRFKNTTA---QIMKPFYVGLPLTDHQSOSVASTQLSALSPALISPLTRAGFTGSLILIRFYSHTV--

680 VPKS-GILMDCKLFFRVAG-----NTAYQLRDRVIDTPEGQDINDI-----CYGCLRQAG

510 VCGVIFILITNNPLGASATGTGTCSPGASSPLVSTHEMKAAGAGRTVLILLEGVLSXATYWGFLBAMX

730 CDHVLNSKARRDKCCVCGGDSCKTACTGTENTYHYGYNTYVLRPAG-ATNIDVROHSFSETDDDNVLLAS

580 HGTVLVS--FWDIILVTLPLIDHVVTVDT--HMLPGCP--LPTGRARAPVQGLPLTGCRTFLPLCHW

800 SSKGEFLNGFNVTVMKARELRIGNAVVEYSG-----SETA-VERINSTDRIEQLLOLVSYGKLYNDVRY

SHQG---STGTLV---SHRALISADVSHGTPPTSGXAGPBRKS-----PLGSWCTV--CASREBG

870 S--FNIPIEDKPOQFYWNS-----HGPMQA--CSKPGOGERKKRLVCTRESDQLTVSDQKDRLP

710 SWGRMTXWAFVTLGISWSSCCITVIVYHXHISPDARHCS-----XIVSLSTSD--TALSPECTRLP

930 OPGHTEPCGTCDDLRMHVARSSECSAOCGLGRTLD-ITYCAKYSRLDGKTEKYDDGFCSSHP-----KP

770 GCTFRLTSLHRCWGLRLVTHSIIISIFTKL--TPDLAVLMPSTGQALXPV-----SRPILMTRELLAP

990 SNREKCSGECNTGGMRY--SAMTECSKCDG-----GTQRRAICVMTRVNDVDDSCSTHOEYVTLQRCSE

830 LN-GNELLMCAEPAVIOYIISIDTNSPSPGLPTVAFYTFESGRAPAIR-----VT--FPAKA

1060 FPCPOKMSGDNSECIVTCKGKHKHQQWQCGEDRLNBMDOPETRKPRISMOTCOQPECASMO--AGPMQCS

880 FETVAV--FGMAAAKTIINLSLPIQPAIFG---TVDOQC-AVTQATILGTTFTAGCMPPQVYVCTGFSN

1130 V-TCGQ--GYQLRAVKCI---IGTYKSVVDNDNCNAATRP-TDQDCEL---PSCHPPPAPAEETRSTYSA

950 VSHLGGPIASLRNSKILRFPGANKFSEPPPLAGFTACLPMTVYPVKLLRETFMNIERLAIVGVLSHR

1190 PRTOMRFGSMTPC---SATCGKGTMRMYVS-----CRDENGSAVDESACATLRPPVAKEEC-----SVT

1020 QNLKQFLLNAICXVNSYFSGLSGPTVLYHSITRNANPFGHCNKKYSIXOEFETLAXSXVYVYCFPYETVL

1240 PCQOMKALDMSSCSVYCGGGRATROVMCVNSDHYIDRESCDODIYPERDQCSMSP--CPQRTDPSGL---

1090 PHINIGSTWNS-----DHSIVTITMYCIKCSCH-----CAXKRIATTNPTFTSPGYVNMILASSLPK

1310 AQHPFONEDYRBRASPSRTIHLGDNQWRTGPMGACSSCTCAGSQRHVCOCDENGYANDCYERIK--PDE

1150 ALDITICVLAHRS---SIYHSYKILVGCVP---CHS---ABQLAPVLRH--NSTYFRDPAHIGKQKTVDV

1380 QRACEGSPCQMAYGWMBCTKLGCGG---IRTRLVYCCORSNGERPRDLSCEILLKPPRREOCNTHACPHDA

1210 KMLPVAVSTLFISEVSLLETWLRVAGLRFHSTSTYVSTTFEMSVALSNGCFDAP-----STGCSG

Init. Opt.

7

OHSEFSEDDDDNYLALSSSGEFLNGFNVTYMAKREIRIGNAVEYSSSEFAVERIN
 STDRIOELLQVLSVGLKYNPDVRSFNIPIEDRPOQRYNASHPMWQACPCGGER
 KRLVCTRESDDLVSDQCDRLPQPHITTEPCGDCDLRMHVAHSCSACGCGYR
 TLDIYCAKSRDLGTEKYDGDGSHPKPSNRKSCGCTGDMGMRYSMTSCSCD
 GGTORRRACVMTNRNDVLDSCCTHKEKTYIORSCEFPPOKSGDMSBCLTCKGK
 KHQVWQCGEEDRLNDACDDPERKPSKOTQCPEDASAOAPWQACSVTCGQYOLR
 AYKCIITGYMSVDDNDCAATRPDTODCELPCHPPAPETRRSYSAFRTQMR
 GSWTPCSATCGKTRMYVSCDENGSDVADBSACATLPBPVAKESCVTPCGOMKALD
 WSSCVATCGGRATROVMCVNVDHVIDSECDODYIPETDODGSMSPCPORTPSGL
 AOHFQNEDEYRPRRSASPSRTHVILGNGMTPRSPMGWAGSCSTGAGSGORRYVYODENGT
 ANDCVERIKPDORACESGPPCPOMAYVNGMECTKICSGGIRLRYLCCORSMERPD
 SCETLDKPPDRQCNTHACPHDAAMSTGCPWSSCSGCGHGRORRYVYAMADSHLES
 DYCKHLAKPHGRKCGRCRPMKAGAMQSCVSGRGVQORHVGCCQIGTHIAETE
 CNPTRESEKDCQGRCPCLYTMRAEMQECTKTCGSGSRKRYKVCVCDNKNVEHGAR
 CTVSKRPVRESISLPCEYVITGEMSEVPSMEI."

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 190 Optimized Score = 190 Significance = 2.05
 Residue Identity = 100% Matches = 190 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0
 Translation Frame = 2

NSEHTAVISLCSGMLGTFRRSHDGYFIEPIQSDQEDQEEONKPHIYRSAPOREPSTGRHACDTSEHN
 10 20 30 40 50 60 70

RHSKDKKTRARKWGERINLAGVAAALNSGLATEASVANGKNTDREKTRHRTKRFSTYFREVLEVAD
 80 90 100 110 120 130 140

NRMVSYGENLQHYIITLMSIVASIKDPSISGLNIVNLIVNHEDGSPISNAOTILKRFQOMQSK
 150 160 170 180 190 200 210

NSFGGIHHDPAVLLTRQDICRAHDKDITGLAELGTICDPYRSCSISEDSGLSTATIHAELGHVYNNPHD
 220 230 240 250 260 270 280

160 170 180 X
 NNKKEGVKSPQHYMAPTLNFYTNPMWMSK
 190 200 210 220 230 240 250 260 270 280

KOCLEIFGSGSOVCPYMOCKRMLNANNVGHKGRTOHTPMADGTECEPCHCKYGCVPREMDVPYTDGS
 370 380 390 400 410 420 430

WGSMSPEGTCSTRTGGGITAIRECNRPEPKNGKYCYVRRAKFKSCNTPEPLKOKRDFRDEQCAHEDGKH
 440 450 460 470 480 490 500

NINGLLPVRWPKYSIGILMKDRCKLCFRVAGNTAYYOLRDRVIDGTGCGDPTNDICVGLCRQAGCHVLN
 510 520 530 540 550 560 570

SKARRDKCGVGGDNNSSCKTVAGTFNTVHYGYNTVVRIPAGATNIDVNOHSSGSETDDNDLALSSSGEFL
 580 590 600 610 620 630 640

LNQNFVYMAKREIRIGNAVEYSSSEFAVERINSTDRIOELLQVLSVGLKYNPDVRSFNIPIEDKPOO
 650 660 670 680 690 700 710 720

FTWNSHGPWQACSKPCGGRKRLVCTRESQGLVSDQRCDLRQPGHITTEPCGDCDLRMHVASRSCSCAQ
 730 740 750 760 770 780 790

CGSLGRTLDIYAKSRDLGTEKYDGDGCSHPPKPSNRKSCGECNTGMRYSMTSCSKSCDGTQORRA
 800 810 820 830 840 850 860

ICVNTNRNDVLDSCCTHKEKTYIORSCEFPPOKSGDMSBCLTCKGKHQVWQCGEEDRLNDACDPE
 870 880 890 900 910 920 930
 TPRTSMQTCQOQDECAQWAGPMWQCSVTGQGYOLRAVKCIITGYMSVDDNDCAATRPDTQDCPLPSCH
 940 950 960 970 980 990 1000
 PPPAPETRRSYSAFRTQMRPGSTPESATCGGTRMYVSCDENGSDVADBSACATLPBPVAKESCVTP
 1010 1020 1030 1040 1050 1060 1070 1080
 CGQKALDMSSCSVTCGGRATROYMCVNSDHYIDRECDODYIPETDODGSMSPCPORTPSGLAQHPFO
 1090 1100 1110 1120 1130 1140 1150
 NEDYRPRRSASPSRTHVILGNGMTPRSPMGWAGSCSTGAGSGORRYVYODENGTANDCVERIKPDORACESG
 1160 1170 1180 1190 1200 1210 1220
 CPQWYGNWGECTKICGGGIFTRLVVCCORSNGERPPLDSCETLDKPPDRQCNTHACPHDAAMSTGFWSSCS
 1230 1240 1250 1260 1270 1280 1290
 VSCGRGHKORRYVYAMADSHLES
 1300 1310 1320

2. US-09-972-467-2 (289-478)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_11b: BluescriptIT SK plus
 clone: fh1167.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS Nagase, T., Kikuno, R., Ishikawa, K.I., Hirose, M., and Ohara, O.
 TITLE Prediction of the coding sequences of unidentified human genes.
 JOURNAL XVI. The complete sequences of 150 new cDNA clones from brain which
 MEDLINE code for large proteins in vitro
 20181126
 REFERENCE DNA Res. 7 (1), 65-73 (2000).

2 (bases 1 to 5139)
 Nagase, T., and Kikuno, R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdna@info.kazusa.or.jp)
 URL: http://www.kazusa.or.jp/hunge/, Tel: +81-438-52-3913,
 Fax: +81-438-52-3914

FEATURES
 Source Location/Qualifiers

1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh1167"
 /tissue_type="brain"
 /clone_11b="pluescriptIT SK plus"
 1..4417
 /gene="KIAA1312"
 /gene="KIAA1312"
 /note="Start codon is not identified."

CDS

product="KIAA1312 protein"
 /protein_id="BA92550.1"
 /db_xref="GI:7242978"
 /translation="NSEHTAVISLCSGMLGTFRRSHDGYFIEPIQSDQEDQEEONKPHIYRSAPOREPSTGRHACDTSEHNKPSNRKSCGECNTGMRYSMTSCSKSCDGTQORRA
 LATEFSAVNGKTDNTRKTRHRTKRLSPRVVYVADNRMVSYHGBNLQHYTL
 TLMSTVASTYKDPSTGILNIVNLIVNHEDGSPISFNAOTILKRFQOMQHSKS
 PGGIHDPAVLLTRQDICRAHDKDITGLAELGTICDPYRSCSISEDSGLSTATIHAELGHVYNNPHD

